


```
Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservaive: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.02% Indels: 0
DB: 12 Gaps: 0

US-09-817-318-1 (1-780) x US-10-001-843-134 (1-66)
OY 603 ATGGCGAGGATAAATCAGAACTTCTGTCATACAAAGTTATGTTTATGGGTATTTT 662
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MeClyArqsplyssrsergluValThrValAsnlnlyValMetPheTyGlyTyrPhe 20
OY 663 ATAGTCGATAAATTCATTACT 683
      |||||||||||||||||||
Db 21 lIeGlyAsplysPheIleThr 27

RESULT 2
US-09-925-301-1458
; Sequence 1458, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1458
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (47)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1458

Alignment Scores:
Pred. No.: 68.7 Length: 115
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservaive: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 10 Gaps: 0

US-09-817-318-1 (1-780) x US-09-925-301-1458 (1-115)
OY 584 ATCAAAAGATATTAGATTAAT 604
      |||||||||||||||||||
Db 40 lIeIlysllyStIleuAspAsn 46

RESULT 3
US-09-886-055-327
; Sequence 327, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 327
```

```
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-886-055-327

Alignment Scores:
Pred. No.: 60.9 Length: 291
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservaive: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 10 Gaps: 0

US-09-817-318-1 (1-780) x US-09-886-055-327 (1-291)
OY 637 TTATTGACGACTACTCTGAT 617
      |||||||||||||||||||
Db 46 LeuIeuthrValThrSerAsp 52

RESULT 4
US-09-789-482-4
; Sequence 4, Application US/09789482
; Patent No. US20020054875A1
; GENERAL INFORMATION:
; APPLICANT: Koch, Alisa E.
; APPLICANT: Ruth, Jeffrey H.
; APPLICANT: Rotman, James B.
; TITLE OF INVENTION: Therapeutic Methods That Target
; TITLE OF INVENTION: Fractalkine or CX3CR1
; FILE REFERENCE: 3238,1000-003
; CURRENT APPLICATION NUMBER: US/09/789,482
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: US 60/183,568
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-789-482-4

Alignment Scores:
Pred. No.: 59.3 Length: 355
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservaive: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 10 Gaps: 0

US-09-817-318-1 (1-780) x US-09-789-482-4 (1-355)
OY 704 TTATTCTCGTGTGAATA 724
      |||||||||||||||||||
Db 232 LeuIleuValValIle 238

RESULT 5
US-09-789-486-4
; Sequence 4, Application US/09789486
; Patent No. US20020055456A1
; GENERAL INFORMATION:
; APPLICANT: Koch, Alisa E.
; TITLE OF INVENTION: Therapeutic Methods That Target
; TITLE OF INVENTION: Fractalkine or CX3CR1
; FILE REFERENCE: 3238,1000-004
; CURRENT APPLICATION NUMBER: US/09/789,486
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: US 60/183,568
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 355
```

TYPE: PRT
ORGANISM: Homo sapiens
US-09-789-486-4

Alignment Scores:
Pred. No.: 59.3 Length: 355
Score: 7.00 Matches: 7
Best Local Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 10 Gaps: 0

US-09-817-318-1 (1-780) x US-09-789-486-4 (1-355)

QY 704 TTATTCCTCGTGTGTAATA 724
Db 232 LeuLeuLeuValValIle 238

RESULT 6
US-09-799-777-47

Sequence 47, Application US/09799777
Patent No. US20020091244A1
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
Hillman, Jennifer L.
Corley, Neil C.
Guegler, Karl J.
Baugh, Mariah
Sather, Susan
Shah, Purvi

TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS
NUMBER OF SEQUENCES: 154
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA: US/09/799,777
APPLICATION NUMBER: US/09/799,777
FILING DATE: 06-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/002,485
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BILLINGS, LUCY J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0459 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PANCNOT04
CLONE: 2084489

SEQUENCE DESCRIPTION: SEQ ID NO: 47 :
US-09-799-777-47

Alignment Scores:
Pred. No.: 58.3 Length: 402
Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 10 Gaps: 0

US-09-817-318-1 (1-780) x US-09-799-777-47 (1-402)

QY 760 GTAAAGCGMAAAAAAAAAA 780
Db 223 ValLysAlaLysLysLys 229

RESULT 7
US-10-055-228-5

Sequence 5, Application US/10055228
Publication No. US20030022316A1
GENERAL INFORMATION:
APPLICANT: Fox, Brian A.
APPLICANT: Taft, David W.
APPLICANT: Sheppard, Paul O.
TITLE OF INVENTION: CUB DOMAIN PROTEIN ZCUB3 AND MATERIALS
FILE REFERENCE: 00-60
CURRENT APPLICATION NUMBER: US/10/055,228
CURRENT FILING DATE: 2002-01-23
PRIOR APPLICATION NUMBER: 60/263,989
PRIOR FILING DATE: 2001-01-24
NUMBER OF SEQ ID NOS: 23
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 419
TYPE: PRT
ORGANISM: Homo sapiens
US-10-055-228-5

Alignment Scores:
Pred. No.: 58 Length: 419
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 9 Gaps: 0

US-09-817-318-1 (1-780) x US-10-055-228-5 (1-419)

QY 404 TTGTTAACTAGTTTCTTCG 424
Db 10 LeuLeuThrSerPheLeuLeu 16

RESULT 8
US-10-055-228-3

Sequence 3, Application US/10055228
Publication No. US20030022316A1
GENERAL INFORMATION:
APPLICANT: Fox, Brian A.
APPLICANT: Taft, David W.
APPLICANT: Sheppard, Paul O.
TITLE OF INVENTION: CUB DOMAIN PROTEIN ZCUB3 AND MATERIALS
FILE REFERENCE: 00-60
CURRENT APPLICATION NUMBER: US/10/055,228
CURRENT FILING DATE: 2002-01-23
PRIOR APPLICATION NUMBER: 60/263,989
PRIOR FILING DATE: 2001-01-24
NUMBER OF SEQ ID NOS: 23
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 444
TYPE: PRT
ORGANISM: Homo sapiens
US-10-055-228-3

Alignment Scores:
Pred. No.: 57.6 Length: 444

Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 9 Gaps: 0

US-09-817-318-1 (1-780) x US-10-055-228-3 (1-444)

QY 404 TTGTTACTAGTTTCTCTG 424
|||||
Db 10 LeuLeuThrSerPheLeuLeu 16

RESULT 9

US-09-910-186A-14
; Sequence 14, Application US/09910186A
; Publication No. US2003009025A1
; GENERAL INFORMATION:
; APPLICANT: U.S. Army Medical Research & Material Command
; TITLE OF INVENTION: RECOMBINANT VACCINE AGAINST BOTULINUM
; FILE REFERENCE: A33626-A 067252.0107
; CURRENT APPLICATION NUMBER: US/09/910.186A
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: PCT/US00/12890
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: 09/611,419
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 60/133,865
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,866
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,867
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,868
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,869
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,873
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 08/123,975
; PRIOR FILING DATE: 1993-09-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-09-910-186A-14

Alignment Scores:
Pred. No.: 57.5 Length: 449
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 9 Gaps: 0

US-09-817-318-1 (1-780) x US-09-910-186A-14 (1-449)

QY 607 GCAGGATAATCGAAGTGA 627
|||||
Db 164 AlaGlyIleAsnGlnLysLeu 170

RESULT 10

US-09-098-079-22
; Sequence 22, Application US/09098079
; Patent No. US20020064773A1
; GENERAL INFORMATION:
; APPLICANT: Herznstadt, Corina
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Cleveneger, William

APPLICANT: Fahy, Eoin F.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: DIAGNOSTIC METHOD BASED ON QUANTIFICATION OF
; TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/098,079
FILING DATE: 15-JUN-1998
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Roseman Ph.D., Stephen J.

REGISTRATION NUMBER: 43,058

REFERENCE/DOCKET NUMBER: 660088.416

TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

US-09-098-079-22

Alignment Scores:
Pred. No.: 57.3 Length: 459
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 10 Gaps: 0

US-09-817-318-1 (1-780) x US-09-098-079-22 (1-459)

QY 649 AACATACTTGTTATGACA 629
|||||
Db 390 AsnIleThrLeuLeuThr 396

RESULT 11

US-09-925-297-604
; Sequence 604, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 604
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (18)


```
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (25)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (57)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (551)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-604

Alignment Scores:
Pred. No.: 55.4      Length: 595
Score: 7.00         Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92%   Indels: 0
DB: 10              Gaps: 0

US-09-817-318-1 (1-780) x US-09-925-297-604 (1-595)

QY 88 TATGCTACAGCTTTCATCT 68
   |||||
Db 584 TyrsertyserPheSerSer 590

RESULT 12
US-09-764-864-1282
; Sequence 1282, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1282
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1282

Alignment Scores:
Pred. No.: 55.4      Length: 600
Score: 7.00         Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92%   Indels: 0
DB: 10              Gaps: 0

US-09-817-318-1 (1-780) x US-09-764-864-1282 (1-600)

QY 303 TTGTAAACGAGACCTCTTT 283
   |||||
Db 22 PheValAsnArgThrLeuphe 28

RESULT 13
US-09-925-301-1094
; Sequence 1094, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
```

```
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1094
; LENGTH: 615
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (2)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (21)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (113)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (132)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (155)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (156)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (157)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1094

Alignment Scores:
Pred. No.: 55.2      Length: 615
Score: 7.00         Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92%   Indels: 0
DB: 10              Gaps: 0

US-09-817-318-1 (1-780) x US-09-925-301-1094 (1-615)

QY 303 TTGTAAACGAGACCTCTTT 283
   |||||
Db 37 PheValAsnArgThrLeuphe 43

RESULT 14
US-09-771-161A-209
; Sequence 209, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; PRIORITY FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 209
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-09-771-161A-209

Alignment Scores:

Pred. No.:	54.5	Length:	676
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.86%	Indels:	0
DB:	10	Gaps:	0

US-09-817-318-1 (1-780) x US-09-771-161A-209 (1-676)

QY 610 GGGATTAATCAGAGTTACTG 630
|||||

DB 281 GYIYleasnglnhlsleu 287

RESULT 15

US-09-764-864-823

; Sequence 823, Application US/09764864

; Patent No. US20020132753A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PT23

; CURRENT APPLICATION NUMBER: US/09/764,864

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 1792

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 823

; LENGTH: 759

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (18)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (19)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (21)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (257)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (299)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-764-864-823

Alignment Scores:

Pred. No.:	53.7	Length:	759
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	10	Gaps:	0

US-09-817-318-1 (1-780) x US-09-764-864-823 (1-759)

QY 303 TTTGTAACAGAGCCTCTTT 283
|||||

DB 181 PhevalasnaargThrlauphe 187

RESULT 16

US-09-919-408-2

; Sequence 2, Application US/09919408

; Patent No. US20020072077A1

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Thor R.

; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL RECEPTORS AND THEIR LIGANDS

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Imclone Systems Incorporated

STREET: 180 Varick Street

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10014

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/919,408

FILING DATE: 31-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/977,451

FILING DATE: <Unknown>

APPLICATION NUMBER: US 07/906,397

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: US PCT/US92/05401

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: TW 81102961

FILING DATE: 15-APR-1992

APPLICATION NUMBER: US PCT/US92/02750

FILING DATE: 02-APR-1992

APPLICATION NUMBER: US 07/813,593

FILING DATE: 24-DEC-1991

APPLICATION NUMBER: US 07/793,065

FILING DATE: 15-NOV-1991

APPLICATION NUMBER: US 07/728,913

FILING DATE: 28-JUN-1991

APPLICATION NUMBER: US 07/679,666

FILING DATE: 02-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Pelt, Irving N.

REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: LEM-3-7P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405

TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 992 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-919-408-2

Alignment Scores:

Pred. No.:	51.9	Length:	992
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.86%	Indels:	0
DB:	10	Gaps:	0

US-09-817-318-1 (1-780) x US-09-919-408-2 (1-992)

QY 251 CTCTTCAATGACTTGAC 271
|||||

DB 245 Leupherthrlspleuasn 251

RESULT 17

US-09-872-136-2

; Sequence 2, Application US/09872136

; Patent No. US20020119545A1

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Thor R.

; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

```

: RECEPTORS AND THEIR LIGANDS
:
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESS: Imclone Systems Incorporated
: STREET: 180 Varick Street
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10014
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/872.136
: FILING DATE: 01-Jun-2001
: CLASSIFICATION: <Unknown>
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/09/208.786
: FILING DATE: <Unknown>
: APPLICATION NUMBER: US/09/021.324
: FILING DATE: <Unknown>
: APPLICATION NUMBER: US/07/977.451
: FILING DATE: 1992-11-19
: APPLICATION NUMBER: US 07/906.397
: FILING DATE: 26-JUN-1992
: APPLICATION NUMBER: US PCT/US92/05401
: FILING DATE: 26-JUN-1992
: APPLICATION NUMBER: TW 81102961
: FILING DATE: 15-APR-1992
: APPLICATION NUMBER: US PCT/US92/02750
: FILING DATE: 02-APR-1992
: APPLICATION NUMBER: US 07/813.593
: FILING DATE: 24-DEC-1991
: APPLICATION NUMBER: US 07/793.065
: FILING DATE: 15-NOV-1991
: APPLICATION NUMBER: US 07/728.913
: FILING DATE: 28-JUN-1991
: APPLICATION NUMBER: US 07/679.666
: FILING DATE: 02-APR-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Feil, Irving N.
: REGISTRATION NUMBER: 28.601
: REFERENCE/DOCKET NUMBER: LEM-3-7P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-645-1405
: TELEFAX: 212-645-2054
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 992 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-872-136-2

Alignment Scores:
Pred. No.: 51.9 Length: 992
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 10 Gaps: 0

US-09-817-318-1 (1-780) x US-09-872-136-2 (1-992)
OY 251 CTCCTCACATTGACTTGAC 271
Db 245 LeuphethrIIeaspleuash 251

RESULT 18
US-09-919-408-4
```

```

: Sequence 4, Application US/09919408
: Patent No. US20020072077A1
:
: GENERAL INFORMATION:
: APPLICANT: Lemischka, Ihor R.
: TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
: RECEPTORS AND THEIR LIGANDS
:
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESS: Imclone Systems Incorporated
: STREET: 180 Varick Street
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10014
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/919.408
: FILING DATE: 31-Jul-2001
: CLASSIFICATION: <Unknown>
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/977.451
: FILING DATE: <Unknown>
: APPLICATION NUMBER: US 07/906.397
: FILING DATE: 26-JUN-1992
: APPLICATION NUMBER: US PCT/US92/05401
: FILING DATE: 26-JUN-1992
: APPLICATION NUMBER: TW 81102961
: FILING DATE: 15-APR-1992
: APPLICATION NUMBER: US PCT/US92/02750
: FILING DATE: 02-APR-1992
: APPLICATION NUMBER: US 07/813.593
: FILING DATE: 24-DEC-1991
: APPLICATION NUMBER: US 07/793.065
: FILING DATE: 15-NOV-1991
: APPLICATION NUMBER: US 07/728.913
: FILING DATE: 28-JUN-1991
: APPLICATION NUMBER: US 07/679.666
: FILING DATE: 02-APR-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Feil, Irving N.
: REGISTRATION NUMBER: 28.601
: REFERENCE/DOCKET NUMBER: LEM-3-7P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-645-1405
: TELEFAX: 212-645-2054
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 993 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-919-408-4

Alignment Scores:
Pred. No.: 51.9 Length: 993
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 10 Gaps: 0

US-09-817-318-1 (1-780) x US-09-919-408-4 (1-993)
OY 251 CTCCTCACATTGACTTGAC 271
Db 244 LeuphethrIIeaspleuash 250

RESULT 19
```

US-09-872-136-4
; Sequence 4, Application US/09872136
; Patent No. US20020119545A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varlick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/872,136
; FILING DATE: 01-Jun-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/208,786
; FILING DATE: <unknown>
; APPLICATION NUMBER: US/09/021,324
; FILING DATE: <unknown>
; APPLICATION NUMBER: US/07/977,451
; FILING DATE: 1992-11-19
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 993 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-872-136-4

Alignment Scores:
Pred. No.: 51.9 Length: 993
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
Caps: 0
US-09-817-318-1 (1-780) x US-09-872-136-4 (1-993)

OY 251 CTCCTGCAATTCGCTTGAC 271
Db 244 Leuphethrilleaspleunsn 250
RESULT 20
US-09-742-096-3
; Sequence 3, Application US/09742096
; Patent No. US20020155441A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERGES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 200773USODIV
; CURRENT APPLICATION NUMBER: US/09/742,096
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 3
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: P. falciparum
US-09-742-096-3

Alignment Scores:
Pred. No.: 48 Length: 1786
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
Gaps: 0
US-09-817-318-1 (1-780) x US-09-742-096-3 (1-1786)
OY 760 GTAAGGCCAATAAAAAAAAAA 780
Db 1583 VALLYSALALYSLSYLSYLS 1595
RESULT 21
US-09-802-669-46
; Sequence 46, Application US/09802669
; Patent No. US20020004490A1
; GENERAL INFORMATION:
; APPLICANT: Dean, Nicholas M.
; APPLICANT: Marcussen, Eric G.
; APPLICANT: Wyatt, Jacqueline
; APPLICANT: Zhang, Hong
; TITLE OF INVENTION: Antisense Compound Modulation of Fas Mediated Signaling
; FILE REFERENCE: ISPH-545
; CURRENT APPLICATION NUMBER: US/09/802,669
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 09/665,615
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 09/290,640
; PRIOR FILING DATE: 1999-04-12
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 46
; LENGTH: 2485
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-802-669-46

Alignment Scores:
Pred. No.: 46 Length: 2485
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.92% Indels: 0
DB: 10 Gaps: 0
US-09-817-318-1 (1-780) x US-09-802-669-46 (1-2485)
QY 503 CTTTCAAAAGCATGGCTTT 483
|||||
Db 237 LeuserlysserMetGlypne 243
RESULT 22
US-09-927-436-10
; Sequence 10, Application US/09927436
; Patent No. US20020155455A1
; GENERAL INFORMATION:
; APPLICANT: Tadayoni-Rebek, Miltra
; APPLICANT: Amshay, Joseph W.
; APPLICANT: Rooney, Regina
; TITLE OF INVENTION: Highly Homogeneous Molecular Markers for Electrophoresis
; FILE REFERENCE: 0942.5300001
; CURRENT APPLICATION NUMBER: US/09/927,436
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: US 60/224,345
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Synthetic peptide
; NAME/KEY: MOD_RES
; LOCATION: (3)..(3)
; OTHER INFORMATION: Modified with TMR
; NAME/KEY: MOD_RES
; LOCATION: (14)..(14)
; OTHER INFORMATION: Modified with TMR
; NAME/KEY: MOD_RES
; LOCATION: (15)..(15)
; OTHER INFORMATION: AMIDATION
US-09-927-436-10
Alignment Scores:
Pred. No.: 870 Length: 15
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.45% Indels: 0
DB: 9 Gaps: 0
US-09-817-318-1 (1-780) x US-09-927-436-10 (1-15)
QY 763 AAGGCAAAAAAAAAAAAA 780
|||||
Db 8 Lysalalyslyslslys 13
RESULT 23
US-09-805-301-84
; Sequence 84, Application US/09805301
; Patent No. US20020173456A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Louis C.
; APPLICANT: Sparrow, James T.
; APPLICANT: Hauer, Jochen
; APPLICANT: Mims, Martha P.
; TITLE OF INVENTION: LIPOPHILIC PEPTIDES FOR
; MACROMOLECULE DELIVERY
; NUMBER OF SEQUENCES: 139
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700

CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 6.0
SOFTWARE: Word Perfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/805,301
FILING DATE: 12-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/584,043
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 217/189
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 84:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 84:
US-09-805-301-84
Alignment Scores:
Pred. No.: 863 Length: 16
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.45% Indels: 0
DB: 9 Gaps: 0
US-09-817-318-1 (1-780) x US-09-805-301-84 (1-16)
QY 763 AAGGCAAAAAAAAAAAAA 780
|||||
Db 1 Lysalalyslyslslys 6
RESULT 24
US-09-826-752-24
; Sequence 24, Application US/09826752
; Patent No. US20010026930A1
; GENERAL INFORMATION:
; APPLICANT: Guarente, Leonard P.
; APPLICANT: Austriaco Jr., Nicanor
; APPLICANT: Claus, James J.
; APPLICANT: Cole, Francesca
; APPLICANT: Kennedy, Brian
; TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESCENCE IN
; YEAST
; FILE REFERENCE: 0050.1491-005
; CURRENT APPLICATION NUMBER: US/09/826,752
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 08/396,001
; PRIOR FILING DATE: 1995-02-28
; PRIOR APPLICATION NUMBER: PCT/US94/09351
; PRIOR FILING DATE: 1994-08-15
; PRIOR APPLICATION NUMBER: US 08/107,408
; PRIOR FILING DATE: 1993-08-16
; PRIOR APPLICATION NUMBER: US 09/323,433
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 48

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-826-752-24

Alignment Scores:
Pred. No.:      863      Length:      16
Score:          6.00     Matches:      6
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:     2.45%      Indels:      0
DB:              10        Gaps:         0

US-09-817-318-1 (1-780) x US-09-826-752-24 (1-16)

OY 236 GCTTACGACCTTACTC 253
    |||||||
DB 10 AlaleuGlnTrtleu 15

RESULT 25
; Sequence 38407, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aomic-a-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
```

```
; SEQ ID NO 38407
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004468.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
US-09-864-761-38407

Alignment Scores:
Pred. No.:      850      Length:      18
Score:          6.00     Matches:      6
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:     2.45%      Indels:      0
DB:              10        Gaps:         0

US-09-817-318-1 (1-780) x US-09-864-761-38407 (1-18)

OY 388 CCTTATCCAACTCAWTT 405
    |||||||
DB 3 ProleuserAserphe 8

RESULT 26
US-09-984-245-262
; Sequence 262, Application US/09984245
; Patent No. US20020165374A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/984,245
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
```

```
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 262
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-245-262

Alignment Scores:
Pred. No.: 838          Length: 20
Score: 6.00            Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.50%      Indels: 0
DB: 9                  Gaps: 0

US-09-817-318-1 (1-780) x US-09-984-245-262 (1-20)
OY 742 GAGTCTCGAGTCACCA 725
Db 2 GluserProGluserPro 7

RESULT 27
US-09-864-761-37759
; Sequence 37759, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 37759
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL132642.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 8.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
US-09-864-761-37759

Alignment Scores:
Pred. No.: 833          Length: 21
Score: 6.00            Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.45%      Indels: 0
DB: 10                  Gaps: 0

US-09-817-318-1 (1-780) x US-09-864-761-37759 (1-21)
OY 415 TTTTCTGCTGAGTCC 432
Db 11 PheServerValseiser 16

RESULT 28
US-09-864-761-45813
; Sequence 45813, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2001-01-30
```

```

; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45813
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL138807.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.65
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.51
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.64
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.83
; OTHER INFORMATION: EST_HUMAN HIT: BF364684.1, EVALUATE 1.20e+00
US-09-864-761-45813

Alignment Scores:
Pred. No.: 814 Length: 25
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.50% Indels: 0
Gaps: 0
DB: 10

US-09-817-318-1 (1-780) x US-09-864-761-45813 (1-25)
QY 568 TCACATCTAGCATACATA 551
DB 3 SerTherSerSerThr 8

RESULT 29
; Sequence 45698, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemlica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
```

```

; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45698
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC023134.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.75
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.81
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.66
US-09-864-761-45698

Alignment Scores:
Pred. No.: 806 Length: 27
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.45% Indels: 0
Gaps: 0
DB: 10

US-09-817-318-1 (1-780) x US-09-864-761-45698 (1-27)
QY 667 GTGATAATTCATCTACTG 684
DB 15 VallieasSerLeuLeu 20

RESULT 30
; Sequence 6, Application US/08908884
; Patent No. US20020138672A1
; GENERAL INFORMATION:
; APPLICANT: Dong et al.
; TITLE OF INVENTION: ACQUIRED RESISTANCE GENES AND USES THEREOF
; NUMBER OF SEQUENCES: 28
```



```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/908,884
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/023,851
FILING DATE: August 9, 1996
APPLICATION NUMBER: 60/035,166
FILING DATE: January 10, 1997
APPLICATION NUMBER: 60/046,769
FILING DATE: May 16, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/339004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-908-884-6

Alignment Scores:
Pred. No.: 802 Length: 28
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.50% Indels: 0
DB: 8 Gaps: 0

US-09-817-318-1 (1-780) x US-08-908-884-6 (1-28)

QY 159 TTAGTGAAGTTGCTATTA 142
|||||
Db 20 LeuValIysIeuLeu 25

RESULT 31
US-09-864-761-41603
Sequence 41603, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
```

```

PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 41603
LENGTH: 28
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC008506.4
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 20
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 22
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 17
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 16
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 20
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 32
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 20
US-09-864-761-41603

Alignment Scores:
Pred. No.: 802 Length: 28
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.45% Indels: 0
DB: 10 Gaps: 0

US-09-817-318-1 (1-780) x US-09-864-761-41603 (1-28)

QY 233 ACAGCTTTACAGACCTTA 250
|||||
Db 16 ThrAlaLeuGlnThrIeu 21

RESULT 32
US-09-908-323-6
Sequence 6, Application US/09908323
Patent No. US20020073447A1
GENERAL INFORMATION:
APPLICANT: Dong et al.
TITLE OF INVENTION: ACQUIRED RESISTANCE GENES AND USES THEREOF
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
```

```
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/908,323
FILING DATE: 17-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/908,884
FILING DATE: <Unknown>
APPLICATION NUMBER: 60/035,166
FILING DATE: January 10, 1997
APPLICATION NUMBER: 60/046,769
FILING DATE: May 16, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/339004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-908-323-6

Alignment Scores:
Pred. No.: 802 Length: 28
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservat: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.50% Indels: 0
DB: 10 Gaps: 0

US-09-817-318-1 (1-780) x US-09-908-323-6 (1-28)
QY 159 TTAGTGAAGTTCCTATTATTA 142
DB 20 LeuValLysLeuLeuLeu 25

RESULT 33
US-09-864-761-45704
Sequence 45704, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeonica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
```

```
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO: 45704
LENGTH: 32
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC013611.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.58
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.52
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.49
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.47
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.5
OTHER INFORMATION: EST_HUMAN HIT: AM812919.1, EVALU8 7.60e+00
US-09-864-761-45704

Alignment Scores:
Pred. No.: 788 Length: 32
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservat: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.50% Indels: 0
DB: 10 Gaps: 0

US-09-817-318-1 (1-780) x US-09-864-761-45704 (1-32)
QY 744 TTGAGTCTCCAGAGTCAC 727
DB 4 LeuSerLeuGlnSerHis 9

RESULT 34
US-09-908-864-8
Sequence 8, Application US/08908884
Patent No. US2002013872A1
GENERAL INFORMATION:
APPLICANT: Dong et al.
TITLE OF INVENTION: ACQUIRED RESISTANCE GENES AND USES THEREOF
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
```

```

1  COMPUTER: IBM Compatible
2  OPERATING SYSTEM: DOS
3  SOFTWARE: FASTSEQ for Windows Version 2.0
4  CURRENT APPLICATION DATA:
5  APPLICATION NUMBER: US/08/908,884
6  FILING DATE:
7  CLASSIFICATION: 800
8  PRIOR APPLICATION DATA:
9  APPLICATION NUMBER: 60/023,851
10 FILING DATE: August 9, 1996
11 APPLICATION NUMBER: 60/035,166
12 FILING DATE: January 10, 1997
13 APPLICATION NUMBER: 60/046,769
14 FILING DATE: May 16, 1997
15 ATTORNEY/AGENT INFORMATION:
16 NAME: Elbing, Karen L.
17 REGISTRATION NUMBER: 35,238
18 REFERENCE/DOCKET NUMBER: 00786/339004
19 TELECOMMUNICATION INFORMATION:
20 TELEPHONE: 617-428-0200
21 TELEFAX: 617-428-7045
22 INFORMATION FOR SEQ ID NO: 8:
23 SEQUENCE CHARACTERISTICS:
24 LENGTH: 33 amino acids
25 TYPE: amino acid
26 STRANDEDNESS: single
27 TOPOLOGY: linear
28 US-08-908-884-8

```

Alignment Scores:		
Pred. No.:	785	3
Score:	6.00	6
Percent Similarity:	100.00%	0
Best Local Similarity:	100.00%	0
Query Match:	2.50%	0
DB:	8	0
		0
		0

US-09-817-318-1 (1-780) x US-08-908-884-8 (1-33)

OY	159	TTAGTGAAGTTCCTATTA	142
Db	17	LeuValLysLeuLeu	22

```

RESULT 35
US-09-864-761-38783
: Sequence 38783, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME- DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Aemica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30

```

```

PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomamax Sequence Listing Engine vers. 1.1
SEQ ID NO: 38783
LENGTH: 33
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL031588.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.4
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.4
US-09-864-761-38783

```

Alignment Scores:	
Pred. No.:	785
Score:	6.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	2.50%
DB:	10
Length:	3
Matches:	6
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-817-318-1 (1-780) x US-09-864-761-38783 (1-33)

Qy	778	TTTTTTTTTGGCTTA	761
Db	1	PhePhePheLeuProLeu	6

RESULT 36
 US-09-908-323-8
 : Sequence 8, Application US/09908323
 : Patent No. US20020073447A1
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Dong et al.
 :
 : TITLE OF INVENTION: ACQUIRED RESISTANCE GENES AND USES THEREOF
 :
 : NUMBER OF SEQUENCES: 28
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Clark & Elbing LLP
 : STREET: 176 Federal Street
 : CITY: Boston
 : STATE: MA
 : COUNTRY: USA
 : ZIP: 02110
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FASTSED for Windows Version 2.0
 :
 : CURRENT APPLICATION DATA:

```

      APPLICATION NUMBER: US/09/908,323
      FILING DATE: 17-Jul-2001
      CLASSIFICATION: <Unknown>
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: 08/908,884
      FILING DATE: <Unknown>
      APPLICATION NUMBER: 60/035,166
      FILING DATE: January 10, 1997
      APPLICATION NUMBER: 60/046,769
      FILING DATE: May 16, 1997
      ATTORNEY/AGENT INFORMATION:
      NAME: Elbing, Karen L.
      REGISTRATION NUMBER: 35,238
      REFERENCE/DOCKET NUMBER: 00786/339004
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: 617-428-0200
      TELEFAX: 617-428-7045
      INFORMATION FOR SEQ ID NO: 8:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 33 amino acids
      TYPE: amino acid
      STRANDEDNESS: single
      TOPOLOGY: linear
      SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-908-323-8

Alignment Scores:
Pred. No.: 785
Score: 6.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.50%
DB: 10
Gaps: 0

US-09-817-318-1 (1-780) x US-09-908-323-8 (1-33)
QY 159 TTGTGAGTGTGCTATTGA 142
Db 17 LeuValLysLeuLeuLeu 22

RESULT 37
US-09-975-143-24
; Sequence 24, Application US/09975143
; Patent No. US20020155513A1
; GENERAL INFORMATION:
; APPLICANT: HSU, Daniel, K.
; APPLICANT: LIU, Fu-tong
; TITLE OF INVENTION: DOWNLING, Christopher, A.
; TITLE OF INVENTION: GALECTIN EXPRESSION IS INDUCED IN
; FILE REFERENCE: DANHSU.001C1
; CURRENT APPLICATION NUMBER: US/09/975,143
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08561
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 34
; TYPE: PRT
; ORGANISM: nematode
US-09-975-143-24

Alignment Scores:
Pred. No.: 782
Score: 6.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.45%
DB: 9
Gaps: 0

US-09-817-318-1 (1-780) x US-09-975-143-24 (1-34)
QY 379 CATCGCTTTCCTTATCC 396
Db 7 HisArgLeuProLeuSer 12

RESULT 38
US-09-729-674-26
; Sequence 26, Application US/09729674
; Patent No. US20010039335A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Werberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; TITLE OF INVENTION: GENETICS INSTITUTE, INC.
; FILE REFERENCE: 6055-64X
; CURRENT APPLICATION NUMBER: US/09/729,674
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-729-674-26

Alignment Scores:
Pred. No.: 771
Score: 6.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.50%
DB: 10
Gaps: 0

US-09-817-318-1 (1-780) x US-09-729-674-26 (1-38)
QY 774 TTTTGTGCTTACAC 757
Db 12 PhePheCysLeuTyrThr 17

RESULT 39
US-09-865-553-4
; Sequence 4, Application US/09865553
; Patent No. US20020055174A1
; GENERAL INFORMATION:
; APPLICANT: Riltner, Karola
; APPLICANT: Jacobs, Eric
; TITLE OF INVENTION: Complex for Transferring an Anionic Substance of Interest
; FILE REFERENCE: 032751-050
; CURRENT APPLICATION NUMBER: US/09/865,553
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/246,083
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 60/277,982
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: EP 00440162.6
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: EP 01440049.3
; PRIOR FILING DATE: 2001-02-27
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
```

```

; SEQ ID NO 4
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: JTS-1-K13
US-09-865-553-4

Alignment Scores:
Pred. No.: 766 Length: 40
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.45% Indels: 0
DB: 10 Gaps: 0

US-09-817-318-1 (1-780) x US-09-865-553-4 (1-40)
QY 763 AAGCAAAAAAAAAAAAAA 780
Db 24 LysAlaLysLysLysLys 29

RESULT 40
US-09-925-301-1680
; Sequence 1680, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1680
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1680

Alignment Scores:
Pred. No.: 763 Length: 41
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.45% Indels: 0
DB: 10 Gaps: 0

US-09-817-318-1 (1-780) x US-09-925-301-1680 (1-41)
QY 279 TCACAAAGCGCTCTGT 296
Db 5 SerGlnArgGlySerCys 10

RESULT 41
US-09-864-761-42005
; Sequence 42005, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
```

```

; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42005
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011989.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 7.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.6
; OTHER INFORMATION: EST_HUMAN HIT: AW52219.1, EVALUATE 2.00e-19
; OTHER INFORMATION: SWISSPROT HIT: P02775, EVALUATE 1.00e-20
US-09-864-761-42005

Alignment Scores:
Pred. No.: 756 Length: 44
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.50% Indels: 0
DB: 10 Gaps: 0

US-09-817-318-1 (1-780) x US-09-864-761-42005 (1-44)
QY 441 ATAGCAAGGAAGTCTAC 424
Db 33 IleGlyLysGlyThrHis 38

RESULT 42
US-09-764-877-1280
; Sequence 1280, Application US/09764877
; Patent No. US20020147140A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 1280
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-877-1280

Alignment Scores:
Pred. No.: 754      Length: 45
Score: 6.00      Matches: 6
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 2.45%      Indels: 0
DB: 10      Gaps: 0

US-09-817-318-1 (1-780) x US-09-764-877-1280 (1-45)
QY 418 TCTTCTGTGAGTTCCTTT 435
      |||||
Db 38 SerSerValSerSerPhe 43

RESULT 43
US-09-873-637-27
; Sequence 27, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 27
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-873-637-27

Alignment Scores:
Pred. No.: 752      Length: 46
Score: 6.00      Matches: 6
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 2.50%      Indels: 0
DB: 10      Gaps: 0

US-09-817-318-1 (1-780) x US-09-873-637-27 (1-46)
QY 447 TCAATTATAGGCAAGGA 430
      |||||
Db 10 SerIleIleGlyLysGly 15

RESULT 44
US-09-993-844-43
; Sequence 43, Application US/09993844
; Patent No. US20020106739A1
; GENERAL INFORMATION:
; APPLICANT: Oakley, Robert H.
; APPLICANT: Barak, Lawrence S.
; APPLICANT: Laporte, Stephane A.
; APPLICANT: Caron, Marc G.
; TITLE OF INVENTION: Modified G-Protein Coupled Receptors
; FILE REFERENCE: 033072-026
```

```
; CURRENT APPLICATION NUMBER: US/09/993,844
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: US 60/245,772
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/260,363
; PRIOR FILING DATE: 2001-01-08
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 43
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-993-844-43

Alignment Scores:
Pred. No.: 744      Length: 50
Score: 6.00      Matches: 6
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 2.45%      Indels: 0
DB: 10      Gaps: 0

US-09-817-318-1 (1-780) x US-09-993-844-43 (1-50)
QY 415 TTTTCTGTGAGTTC 432
      |||||
Db 8 PheSerSerValSerSer 13

RESULT 45
US-09-864-761-45700
; Sequence 45700, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
```

```
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45700
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC021385.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.65
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.48
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.74
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.62
; OTHER INFORMATION: EST_HUMAN HIT: A1932673.1, EVALU6 6.50e+00
US-09-864-761-45700

Alignment Scores:
Pred. No.:      740      Length:      52
Score:          6.00      Matches:      6
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    2.45%      Indels:      0
DB:             10        Gaps:         0

US-09-817-318-1 (1-780) x US-09-864-761-45700 (1-52)
OY  415 TTTTCTCTGTGAGTTC 432
Db   44 PheserSerValSerSer 49
```

Search completed: February 4, 2003, 07:28:17
Job time : 123 secs



-
-
-
-

1

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 4, 2003, 07:17:33 : Search time 23.5 Seconds
(without alignments)
1953.181 Million cell updates/sec

Title: US-09-817-318-1

Perfect score: 245
Sequence: 1 aactactgtctcttaag.....taagcaaaaaaaaaa 780

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 segs, 29422922 residues

Word size: 1

Total number of hits satisfying chosen parameters: 438908

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p_model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09817318/runatc_04022003_071723_6661/app_query.fasta.1.967
-DB=Issued_Patents_AA -OPMT=fastan -SUFFIX=ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=45 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US09817318.ecgn_1.1.12.tunat_04022003_071723_6661 -NCP=6 -ICP=3
-NO_XLXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2.9	7	21	1	US-08-231-730A-31
2	2.9	7	21	1	US-08-427-001C-31
3	2.9	7	21	1	US-08-457-798-31
4	2.9	7	21	1	US-08-457-171-31
5	2.9	7	21	1	US-08-505-486-31
6	2.9	7	21	3	US-08-689-489C-31
7	2.9	7	21	3	US-08-801-028-31
8	2.9	7	21	3	US-09-340-154-31
9	2.9	7	21	4	US-09-222-802A-31
10	2.9	7	21	4	US-09-462-611B-31
11	2.9	7	21	5	PCT-US94-06176-31
12	2.9	7	21	5	PCT-US94-12550-31

13	2.9	7	21	5	PCT-US95-04335-31	Sequence 31, Appl
14	2.9	7	21	5	PCT-US95-04718-31	Sequence 31, Appl
15	2.9	7	21	5	PCT-US95-09338-31	Sequence 31, Appl
16	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
17	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
18	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
19	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
20	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
21	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
22	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
23	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
24	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
25	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
26	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
27	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
28	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
29	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
30	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
31	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
32	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
33	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
34	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
35	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
36	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
37	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
38	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
39	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
40	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
41	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
42	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
43	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
44	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl
45	2.9	7	21	5	PCT-US95-09339-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-08-231-730A-31
Sequence 31, Application US/08231730A
Patent No. 5561107
GENERAL INFORMATION:
APPLICANT: JAYNES, JESSE M.
INVENTOR: JULIAN, GORDON R.
TITLE OF INVENTION: METHOD OF ENHANCING WOUND HEALING BY STIMULATING FIBROBLASTS
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: STEVEN J. HOUTLOUIS
ADDRESS: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
STREET: 200 PARK DRIVE, SUITE 210
CITY: P.O. BOX 14329
STATE: RESEARCH TRIANGLE PARK
COUNTRY: NORTH CAROLINA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
OPERATING SYSTEM: MACINTOSH
SOFTWARE: M.S. WORD 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/231,730A
FILING DATE: 04-20-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/225,476
FILING DATE: 04-08-94
APPLICATION NUMBER: 08/039,620
FILING DATE: 06-04-93
APPLICATION NUMBER: 08/148,491
FILING DATE: 11-08-93
APPLICATION NUMBER: 08/148,889
FILING DATE: 11-08-93
ATTORNEY/AGENT INFORMATION:

NAME: HULTQUIST, STEVEN J.
REGISTRATION NUMBER: 28021
REFERENCE/DOCKET NUMBER: 4013-106
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)990-9531
TELEFAX: (919)990-9532
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 21
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: PEPTIDE
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-231-730A-31

Alignment Scores:
Pred. No.: 107 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x US-08-231-730A-31 (1-21)

OY 760 GTAAAGCAAAAAAAAAAAAAA 780

Db 15 VallysAlaLysLysLysLys 21

RESULT 2

US-08-427-001C-31

Sequence 31, Application US/08427001C
Patent No. 5717064

GENERAL INFORMATION:
APPLICANT: JULIAN, GORDON R.
TITLE OF INVENTION: METHYLATED LYSINE-RICH LYTIC PEPTIDES.
TITLE OF INVENTION: AND METHOD OF MAKING THE SAME BY REDUCTIVE ALKYLATION
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
STREET: 555 Thirteenth Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/427,001C
FILING DATE: 24-APR-95
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/148,889
FILING DATE: 08-NOV-93
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: WALKER, BARBARA W.
REGISTRATION NUMBER: 35,400
REFERENCE/DOCKET NUMBER: 2093-105A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
TELEFAX: (202)783-6031
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 21

TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
US-08-427-001C-31

Alignment Scores:
Pred. No.: 107 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x US-08-427-001C-31 (1-21)

OY 760 GTAAAGCAAAAAAAAAAAAAA 780

Db 15 VallysAlaLysLysLysLys 21

RESULT 3

US-08-457-798-31

Sequence 31, Application US/08457798
Patent No. 5744445

GENERAL INFORMATION:
APPLICANT: JAYNES, JESSE M.
APPLICANT: JULIAN, GORDON R.
TITLE OF INVENTION: METHOD OF TREATING PULMONARY DISEASE
TITLE OF INVENTION: STATES WITH NON-NATURALLY OCCURRING
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: JESSE M. JAYNES,
ADDRESSEE: DEMETER BIOTECHNOLOGIES, LTD.
STREET: 150 FAYETTEVILLE ST. MALL, SUITE 2700
CITY: RALEIGH
STATE: NORTH CAROLINA
COUNTRY: USA

ZIP: 27601

COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
COMPUTER: APPLE MACINTOSH
OPERATING SYSTEM: MACINTOSH
SOFTWARE: M.S. WORD 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,798
FILING DATE:
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/039,620A

FILING DATE: 19930604

ATTORNEY/AGENT INFORMATION:
NAME: HULTQUIST, STEVEN J.
REGISTRATION NUMBER: 28021
REFERENCE/DOCKET NUMBER: 4013-103
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)990-9531
TELEFAX: (919)990-9532

INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 21
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: PEPTIDE
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-457-798-31

Alignment Scores:
Pred. No.: 107 Length: 21

Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x US-08-457-798-31 (1-21)

QY 760 GTAAAGCGCAAAAAAAAAAAAAA 780

DB 15 VallysAlatylsLysLysLys 21

RESULT 4

US-08-457-171-31
; Sequence 31, Application US/08457171
; Patent No. 5773413
; GENERAL INFORMATION:
; APPLICANT: JAYNES, JESSE M.
; APPLICANT: JULIAN, GORDON R.
; TITLE OF INVENTION: METHOD OF COMBATTING MAMMALIAN NEOPLASIA, AND LYTIC PEPTIDES
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STEVEN J. HULTQUIST
; ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
; STREET: 200 PARK DRIVE, SUITE 210
; STREET: P.O. BOX 14329
; CITY: RESEARCH TRIANGLE PARK
; STATE: NORTH CAROLINA
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
; COMPUTER: APPLE MACINTOSH
; OPERATING SYSTEM: MACINTOSH
; SOFTWARE: M.S. WORD 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,171
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/225,476A
; FILING DATE: 04-08-94
; APPLICATION NUMBER: 08/039,620
; FILING DATE: 06-04-93
; APPLICATION NUMBER: 08/148,491
; FILING DATE: 11-08-93
; APPLICATION NUMBER: 08/148,889
; FILING DATE: 11-08-93
; ATTORNEY/AGENT INFORMATION:
; NAME: HULTQUIST, STEVEN J.
; REGISTRATION NUMBER: 28021
; REFERENCE/DOCKET NUMBER: 4013-106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)990-9531
; TELEFAX: (919)990-9532
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; DESCRIPTION: NO
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-457-171-31

Alignment Scores:

Pred. No.: 107 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.86% Indels: 0
DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x US-08-457-171-31 (1-21)

QY 760 GTAAAGCGCAAAAAAAAAAAAAA 780

DB 15 VallysAlatylsLysLysLys 21

RESULT 5

US-08-505-486-31
; Sequence 31, Application US/08505486
; Patent No. 5955573
; GENERAL INFORMATION:
; APPLICANT: Jesse M. Jaynes
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
; TITLE OF INVENTION: METHODS OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
; STREET: 555 Thirteenth Street N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
; COMPUTER: IBM COMPATIBLE
; OPERATING SYSTEM: DOS
; SOFTWARE: WordPerfect 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/505,486
; FILING DATE: 21-JUL-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/279,472
; FILING DATE: 22-JUL-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WALKER, BARBARA W.
; REGISTRATION NUMBER: 35,400
; REFERENCE/DOCKET NUMBER: 2093-117A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)783-6040
; TELEFAX: (202)783-6031
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; DESCRIPTION: NO
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-505-486-31

Alignment Scores:

Pred. No.: 107 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x US-08-505-486-31 (1-21)

QY 760 GTAAAGCGCAAAAAAAAAAAAAA 780

DB 15 VallysAlatylsLysLysLys 21

```
RESULT 6
US-08-689-489C-31
; Sequence 31, Application US/08689489C
; Patent No. 6001805
; GENERAL INFORMATION:
; APPLICANT: Jesse M. Jaynes, Gordon R. Julian
; TITLE OF INVENTION: Method of Enhancing Wound Healing By
; TITLE OF INVENTION: Stimulating Fibro-blast and Keratinocyte Growth In
; TITLE OF INVENTION: Vivo, Utilizing Amphipathic Peptides
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohnwell, P199, Ernst & Kurz
; STREET: 555 13TH STREET
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/689,489C
; FILING DATE: August 12, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/231,730
; FILING DATE: April 20, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/225,476
; FILING DATE: April 8, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,620
; FILING DATE: June 4, 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/148,889
; FILING DATE: No. 6001805ember 8, 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/148,491
; FILING DATE: No. 6001805ember, 8, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark I. Bowditch
; REGISTRATION NUMBER: 40,315
; REFERENCE/DOCKET NUMBER: 2093-120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: linear
US-08-689-489C-31

Alignment Scores:
Pred. No.: 107 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
Gaps: 0
DB: 3

US-09-817-318-1 (1-780) x US-08-689-489C-31 (1-21)
QY 760 GTAAAGCAAAAAAAAAAAAAA 780
DB 15 VallysAlatylsLysLysLys 21
```

```
RESULT 7
US-08-801-028-31
; Sequence 31, Application US/08801028
; Patent No. 6018102
; GENERAL INFORMATION:
; APPLICANT: JOAN GARBARINO
; APPLICANT: JESSE M. JAYNES
; APPLICANT: WILLIAM BELKNAP
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE CONSTRUCTS, PROTEIN PRO
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STEVEN J. HULTQUIST
; ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
; STREET: 200 PARK DRIVE, SUITE 210
; STREET: P.O. BOX 14329
; CITY: RESEARCH TRIANGLE PARK
; STATE: NORTH CAROLINA
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
; COMPUTER: APPLE MACINTOSH
; OPERATING SYSTEM: MACINTOSH
; SOFTWARE: M.S. WORD 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,028
; FILING DATE: 19-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/279,472
; FILING DATE: JULY 22, 1994
; APPLICATION NUMBER: 08/225,476
; FILING DATE: 04-20-94
; APPLICATION NUMBER: 08/225,476
; FILING DATE: 04-08-94
; APPLICATION NUMBER: 08/039,620
; FILING DATE: 06-04-93
; APPLICATION NUMBER: 08/148,491
; FILING DATE: 11-08-93
; APPLICATION NUMBER: 08/148,889
; FILING DATE: 11-08-93
; ATTORNEY/AGENT INFORMATION:
; NAME: WASSERMAN, FRAN S.
; REGISTRATION NUMBER: 34,273
; REFERENCE/DOCKET NUMBER: 4013-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)990-9531
; TELEFAX: (919)990-9532
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; DESCRIPTION: NO
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-801-028-31

Alignment Scores:
Pred. No.: 107 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
Gaps: 0
DB: 3

US-09-817-318-1 (1-780) x US-08-801-028-31 (1-21)
QY 760 GTAAAGCAAAAAAAAAAAAAA 780
DB 15 VallysAlatylsLysLysLys 21
```

Db 15 VallysAlaLysLysLysLys 21

RESULT 8

US-09-340-154-31

; Sequence 31, Application US/09340154

; Patent No. 6084156

; GENERAL INFORMATION:

; APPLICANT: Jesse M. Jaynes

; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE

; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND

; TITLE OF INVENTION: METHODS OF MAKING AND USING SAME

; NUMBER OF SEQUENCES: 98

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ

; STREET: 555 Thirteenth Street N.W.

; CITY: Washington

; STATE: D. C.

; COUNTRY: USA

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE

; COMPUTER: IBM COMPATIBLE

; OPERATING SYSTEM: DOS

; SOFTWARE: WordPerfect 5.1+

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/340,154

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/505,486

; FILING DATE: 21-JUL-1995

; APPLICATION NUMBER: U.S. 08/279,472

; FILING DATE: 22-JUL-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: WALKER, BARBARA W.

; REGISTRATION NUMBER: 35,400

; REFERENCE/DOCKET NUMBER: 2093-117A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)783-6040

; TELEFAX: (202)783-6031

; INFORMATION FOR SEQ ID NO: 31:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 21

; TYPE: AMINO ACID

; TOPOLOGY: LINEAR

; MOLECULE TYPE:

; DESCRIPTION: PEPTIDE

; HYPOTHETICAL: NO

; FRAGMENT TYPE: COMPLETE PEPTIDE

; ORIGINAL SOURCE: SYNTHETIC

; IMMEDIATE SOURCE: SYNTHETIC

; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED

US-09-340-154-31

Alignment Scores:

Pred. No.:	107	Length:	21
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.86%	Indels:	0
DB:	3	Gaps:	0

US-09-817-318-1 (1-780) x US-09-340-154-31 (1-21)

QY 760 GTAAGCGCAAAAAAAAAA 780

|||||

Db 15 VallysAlaLysLysLysLys 21

RESULT 9

US-09-232-802A-31

; Sequence 31, Application US/09232802A

; Patent No. 6191110

; GENERAL INFORMATION:

; APPLICANT: Jesse M. Jaynes, Gordon R. Julian

; TITLE OF INVENTION: Method of Enhancing Wound Healing By

; Stimulating Fibro-blast and Keratinocyte Growth In

; Vivo, Utilizing Amphipathic Peptides

; NUMBER OF SEQUENCES: 46

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ROTHWELL, Figg, Ernst & Manbeck

; STREET: 555 15TH STREET

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/232,802A

; FILING DATE: 19-Jan-1999

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/689,489

; FILING DATE: August 12, 1996

; APPLICATION NUMBER: US 08/231,730

; FILING DATE: April 20, 1994

; APPLICATION NUMBER: US 08/225,476

; FILING DATE: April 8, 1994

; APPLICATION NUMBER: US 08/039,620

; FILING DATE: June 4, 1993

; APPLICATION NUMBER: 08/148,889

; FILING DATE: No. 619110ember 8, 1993

; APPLICATION NUMBER: 08/148,491

; FILING DATE: No. 619110ember, 8, 1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Mark I. Bowditch

; REGISTRATION NUMBER: 40,315

; REFERENCE/DOCKET NUMBER: 2093-142

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-783-6040

; TELEFAX: 202-783-6031

; INFORMATION FOR SEQ ID NO: 31:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 21 amino acids

; TYPE: amino acid

; STRANDEDNESS: <Unknown>

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; FRAGMENT TYPE: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 31:

US-09-232-802A-31

Alignment Scores:

Pred. No.:	107	Length:	21
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.86%	Indels:	0
DB:	4	Gaps:	0

US-09-817-318-1 (1-780) x US-09-232-802A-31 (1-21)

QY 760 GTAAGCGCAAAAAAAAAA 780

|||||

Db 15 VallysAlaLysLysLysLys 21

RESULT 10

US-09-482-611B-31

; Sequence 31, Application US/09482611B

; Patent No. 6448391

; GENERAL INFORMATION:

; APPLICANT: Garbarino, Joan

; APPLICANT: Belknap, William

;; TITLE OF INVENTION: Ubiquitin-Lytic Peptide Fusion Gene Constructs, Protein Products
;; FILE REFERENCE: 2093-149
;; CURRENT APPLICATION NUMBER: US/09/482,611B
;; PRIOR FILING DATE: 2000-01-14
;; PRIOR APPLICATION NUMBER: US 08/801,028
;; PRIOR FILING DATE: 1997-02-19
;; PRIOR APPLICATION NUMBER: US 08/279,472
;; NUMBER OF SEQ ID NOS: 102
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 31
;; LENGTH: 21
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Lytic Peptide
US-09-482-611B-31

Alignment Scores:
Pred. No.: 107
Score: 7.00 Length: 21
Percent Similarity: 100.00% Matches: 7
Best Local Similarity: 100.00% Conservative: 0
Query Match: 2.86% Mismatches: 0
DB: Indels: 0
Gaps: 0

US-09-817-318-1 (1-780) x US-09-482-611B-31 (1-21)

QY 760 GTAAGGCAAAAAAAAAAAAAA 780
Db 15 VallysAlalyslsylslys 21

RESULT 11
PCT-US94-06176-31
; Sequence 31, Application PC/TUS9406176
; GENERAL INFORMATION:
; APPLICANT: DEMETER BIOTECHNOLOGIES
; TITLE OF INVENTION: METHOD OF TREATING PULMONARY DISEASE STATES WITH NON-NATURALLY
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FRAN S. WASSERMAN
; ADDRESS: INTELLECTUAL PROPERTY/
; STREET: 200 PARK DRIVE, SUITES 209 & 210
; CITY: RESEARCH TRIANGLE PARK
; STATE: NORTH CAROLINA
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
; COMPUTER: APPLE MACINTOSH
; OPERATING SYSTEM: MACINTOSH
; SOFTWARE: M.S. WORD 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/06176
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/039,620
; FILING DATE: 04-06-93
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: WASSERMAN, FRAN S.
; REGISTRATION NUMBER: 34,273
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)990-9531
; TELEFAX: (919)990-9532
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: AMINO ACID

;; TOPOLOGY: LINEAR
;; MOLECULE TYPE:
;; DESCRIPTION: PEPTIDE
;; HYPOTHETICAL: NO
;; FRAGMENT TYPE: COMPLETE PEPTIDE
;; ORIGINAL SOURCE: SYNTHETIC
;; IMMEDIATE SOURCE: SYNTHETIC
;; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
PCT-US94-06176-31

Alignment Scores:
Pred. No.: 107
Score: 7.00 Length: 21
Percent Similarity: 100.00% Matches: 7
Best Local Similarity: 100.00% Conservative: 0
Query Match: 2.86% Mismatches: 0
DB: Indels: 0
Gaps: 0

US-09-817-318-1 (1-780) x PCT-US94-06176-31 (1-21)

QY 760 GTAAGGCAAAAAAAAAAAAAA 780
Db 15 VallysAlalyslsylslys 21

RESULT 12
PCT-US94-12550-31
; Sequence 31, Application PC/TUS9412550
; GENERAL INFORMATION:
; APPLICANT: JULIAN, GORDON R.
; TITLE OF INVENTION: METHYLATED LYSINE-RICH LYTC
; TITLE OF INVENTION: PEPTIDES AND METHOD OF
; TITLE OF INVENTION: MAKING SAME BY REDUCTIVE
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STEVEN J. HULTQUIST
; ADDRESS: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
; STREET: 200 PARK DRIVE, SUITE 210
; CITY: RESEARCH TRIANGLE PARK
; STATE: NORTH CAROLINA
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
; COMPUTER: APPLE MACINTOSH
; OPERATING SYSTEM: MACINTOSH
; SOFTWARE: M.S. WORD 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/12550
; FILING DATE: NOVEMBER 8, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: HULTQUIST, STEVEN J.
; REGISTRATION NUMBER: 28021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)990-9531
; TELEFAX: (919)990-9532
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
PCT-US94-12550-31

Alignment Scores:

Pred. No.: 107 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 5 Gaps: 0

US-09-817-318-1 (1-780) x PCT-US94-12550-31 (1-21)

QY 760 GTAAAGCGCAAAAAAAAAA 780

Db 15 VallysAlatylslyslslys 21

RESULT 13

PCT-US95-04335-31

; Sequence 31, Application PC/TUS9504335

; GENERAL INFORMATION:

; APPLICANT: JAYNES, JESSE M.

; APPLICANT: JULIAN, GORDON R.

; TITLE OF INVENTION: METHOD OF COMBATTING MAMMALIAN NEOPLASIA, AND LYTIC PEPTIDES

; NUMBER OF SEQUENCES: 40

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: FRAN S. MASSEMAN

; ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW

; STREET: 200 PARK DRIVE, SUITE 210

; STREET: P.O. BOX 14329

; CITY: RESEARCH TRIANGLE PARK

; STATE: NORTH CAROLINA

; COUNTRY: USA

; ZIP: 27709

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE

; COMPUTER: APPLE MACINTOSH

; OPERATING SYSTEM: MACINTOSH

; SOFTWARE: M.S. WORD 6.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/04335

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: USA08/225,476

; FILING DATE: 04-08-94

; ATTORNEY/AGENT INFORMATION:

; NAME: MASSEMAN, FRAN S.

; REGISTRATION NUMBER: 34273

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (919)990-9531

; TELEFAX: (919)990-9532

; INFORMATION FOR SEQ ID NO: 31:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 21

; TYPE: AMINO ACID

; TOPOLOGY: LINEAR

; MOLECULE TYPE:

; DESCRIPTION: PEPTIDE

; HYPOTHETICAL: NO

; FRAGMENT TYPE: COMPLETE PEPTIDE

; ORIGINAL SOURCE: SYNTHETIC

; IMMEDIATE SOURCE: SYNTHETIC

; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED

; PCT-US95-04335-31

Alignment Scores:

Pred. No.: 107 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 5 Gaps: 0

US-09-817-318-1 (1-780) x PCT-US95-04335-31 (1-21)

QY 760 GTAAAGCGCAAAAAAAAAA 780

Db 15 VallysAlatylslyslslys 21

RESULT 14

PCT-US95-04718-31

; Sequence 31, Application PC/TUS9504718

; GENERAL INFORMATION:

; APPLICANT: DEMETER BIOTECHNOLOGIES, LTD.

; TITLE OF INVENTION: METHOD OF ENHANCING WOUND HEALING BY STIMULATING FIBROBLAST

; NUMBER OF SEQUENCES: 46

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: FRAN S. MASSEMAN

; ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW

; STREET: 200 PARK DRIVE, SUITE 210

; STREET: P.O. BOX 14329

; CITY: RESEARCH TRIANGLE PARK

; STATE: NORTH CAROLINA

; COUNTRY: USA

; ZIP: 27709

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE

; COMPUTER: APPLE MACINTOSH

; OPERATING SYSTEM: MACINTOSH

; SOFTWARE: M.S. WORD 6.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/04718

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/231,730

; FILING DATE: 20-04-94

; ATTORNEY/AGENT INFORMATION:

; NAME: MASSEMAN, FRAN S.

; REGISTRATION NUMBER: 34273

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (919)990-9531

; TELEFAX: (919)990-9532

; INFORMATION FOR SEQ ID NO: 31:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 21

; TYPE: AMINO ACID

; TOPOLOGY: LINEAR

; MOLECULE TYPE:

; DESCRIPTION: PEPTIDE

; HYPOTHETICAL: NO

; FRAGMENT TYPE: COMPLETE PEPTIDE

; ORIGINAL SOURCE: SYNTHETIC

; IMMEDIATE SOURCE: SYNTHETIC

; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED

; PCT-US95-04718-31

Alignment Scores:

Pred. No.: 107 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 5 Gaps: 0

US-09-817-318-1 (1-780) x PCT-US95-04718-31 (1-21)

QY 760 GTAAAGCGCAAAAAAAAAA 780

Db 15 VallysAlatylslyslslys 21

RESULT 15

PCT-US95-09338-31

; Sequence 31, Application PC/TUS9509338

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE

; CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND

IMMEDIATE SOURCE: SYNTHETIC

OTHER INFORMATION: amino acid residue.

OTHER INFORMATION: amino acid residue.

Alignment Scores:

Alignment Scores:
Pred. No.: 99 Length: 38
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x US-08-484-905-85 (1-38)

OY 625 ACTTGATTTATCCCTGCC 605

Db 19 ThrsrAspLeuSerLeuPro 25

RESULT 18

US-08-481-985B-85
; Sequence 85, Application US/08481985B
; Patent No. 6011146
; GENERAL INFORMATION:
; APPLICANT: Motiez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,985B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0106-04000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Region
; LOCATION: 14..15
; OTHER INFORMATION: /note="Xaa is an unidentified
; OTHER INFORMATION: amino acid residue."
US-08-481-985B-85

Alignment Scores:

Alignment Scores:
Pred. No.: 99 Length: 38
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 3 Gaps: 0

US-09-817-318-1 (1-780) x US-08-481-985B-85 (1-38)

OY 625 ACTTGATTTATCCCTGCC 605

Db 19 ThrsrAspLeuSerLeuPro 25

RESULT 19

US-08-370-476-85
; Sequence 85, Application US/08370476
; Patent No. 6153408
; GENERAL INFORMATION:
; APPLICANT: Motiez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; APPLICANT: Lone, Yu-Chun
; APPLICANT: Ojcius, David
; APPLICANT: Carrouge, Amanda
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,476
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/117,575
; FILING DATE: 07-SEP-1993
; APPLICATION NUMBER: US 08/072,787
; FILING DATE: 06-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05243.0001-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Region
; LOCATION: 14..15
; OTHER INFORMATION: /note="Xaa is an unidentified
; OTHER INFORMATION: amino acid residue."
US-08-370-476-85

Alignment Scores:

Alignment Scores:
Pred. No.: 99 Length: 38
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0

Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 4 Gaps: 0

US-09-817-318-1 (1-780) x US-08-370-476-85 (1-38)

OY 625 ACTTGATTTATCCCTGCC 605
|||||

Db 19 ThrsrAsPleuSerLeuPro 25

RESULT 20
US-08-484-905-86
; Sequence 86, Application US/08484905
; Patent No. 5976551
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; TITLE OF INVENTION: An Altered Major Histocompatibility
; TITLE OF INVENTION: Complex(MHC) Determinant and Methods for Using the
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunnet
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,905
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Potler, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 03495.0106-03000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Region
; LOCATION: 18..19
; OTHER INFORMATION:
; OTHER INFORMATION: /note="Xaa is an unidentified
; amino acid residue."
US-08-484-905-86

Alignment Scores:
Pred. No.: 97.7 Length: 42
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.92% Indels: 0
DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x US-08-484-905-86 (1-42)

OY 625 ACTTGATTTATCCCTGCC 605
|||||

Db 23 ThrsrAsPleuSerLeuPro 29

RESULT 21
US-08-481-985B-86
; Sequence 86, Application US/08481985B
; Patent No. 6011146
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunnet
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,985B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0106-04000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Region
; LOCATION: 18..19
; OTHER INFORMATION:
; OTHER INFORMATION: /note="Xaa is an unidentified
; amino acid residue."
US-08-481-985B-86

Alignment Scores:
Pred. No.: 97.7 Length: 42
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 3 Gaps: 0

US-09-817-318-1 (1-780) x US-08-481-985B-86 (1-42)


```
US-09-366-887A-7
Alignment Scores:
Pred. No.: 87.5 Length: 98
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
Gaps: 0
DB: 4

US-09-817-318-1 (1-780) x US-09-366-887A-7 (1-98)

Qy 640 TTGTATTGACAGTACTTCT 620
    |||||
Db 10 LeuLeuLeuThryAlThrSer 16

RESULT 25
US-09-071-035-292
; Sequence 292, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 292:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-292

Alignment Scores:
Pred. No.: 86.8 Length: 105
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
Gaps: 0
DB: 4

US-09-817-318-1 (1-780) x US-09-071-035-292 (1-105)

Qy 616 AATCAGAGTACTGTTCATA 636
    |||||
Db 31 AsnGlnLysLeuLeuSerIle 37

US-09-069-023-14
; Sequence 14, Application US/09069023A
; Patent No. 6348573
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Koseki, Takeyoshi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
; TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
; FILE REFERENCE: UM-03333
; CURRENT APPLICATION NUMBER: US/09/069,023A
; CURRENT FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-069-023-14

Alignment Scores:
Pred. No.: 86.6 Length: 107
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
Gaps: 0
DB: 4

US-09-817-318-1 (1-780) x US-09-069-023-14 (1-107)

Qy 281 ACAAGAGGCTCTGTTTACA 301
    |||||
Db 22 ThrLysArgValLeuPheThr 28

RESULT 27
US-09-071-035-290
; Sequence 290, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 290:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 131 amino acids
; TYPE: amino acid
```

```

: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-09-071-035-290

Alignment Scores:
Pred. No.: 84.3 Length: 131
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 4 Gaps: 0

US-09-817-318-1 (1-780) x US-09-071-035-290 (1-131)

OY 616 AATCAGAGTACTGTCAATA 636
DB 53 AsnGlnLysLeuLeuSerIle 59

RESULT 28
US-09-134-001C-5047
: Sequence 5047, Application US/09134001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
: TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GTC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: PRIOR FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 5047
: LENGTH: 206
: TYPE: PRT
: ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5047

Alignment Scores:
Pred. No.: 79.5 Length: 206
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 4 Gaps: 0

US-09-817-318-1 (1-780) x US-09-134-001C-5047 (1-206)

OY 582 CTCCTAAGTACTGATGACAT 562
DB 61 ValLeuThrAspSerGlnHis 67

RESULT 29
US-09-415-522-2
: Sequence 2, Application US/09415522A
: Patent No. 6291660
: GENERAL INFORMATION:
: APPLICANT: Gaffney, Thomas
: APPLICANT: Wendland, Juergen
: APPLICANT: Philippsen, Peter
: TITLE OF INVENTION: No. 6291660e1 Fungal Genes Required For No. 6291660ma1 Growth And
: TITLE OF INVENTION: Development
: FILE REFERENCE: CGC2046
: CURRENT APPLICATION NUMBER: US/09/415,522A
: PRIOR FILING DATE: 1999-10-08
: NUMBER OF SEQ ID NOS: 28
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 207
: TYPE: PRT
```

```

: ORGANISM: Ashbya gossypii
US-09-415-522-2

Alignment Scores:
Pred. No.: 79.5 Length: 207
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 4 Gaps: 0

US-09-817-318-1 (1-780) x US-09-415-522-2 (1-207)

OY 44 ATTTTACTAGGCTGTAGGCC 24
DB 119 IleLeuValGlyCysIlySala 125

RESULT 30
US-09-069-023-12
: Sequence 12, Application US/09069023A
: Patent No. 6348573
: GENERAL INFORMATION:
: APPLICANT: Nunez, Gabriel
: APPLICANT: Inohara, Naohiro
: APPLICANT: Koseki, Takeyoshi
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
: TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
: FILE REFERENCE: UM-03333
: CURRENT APPLICATION NUMBER: US/09/069,023A
: PRIOR FILING DATE: 1998-04-27
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 12
: LENGTH: 219
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-069-023-12

Alignment Scores:
Pred. No.: 78.9 Length: 219
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 4 Gaps: 0

US-09-817-318-1 (1-780) x US-09-069-023-12 (1-219)

OY 281 ACAAGAGGTCCTGTTTACA 301
DB 22 ThrLysArgValLeuPheThr 28

RESULT 31
US-09-134-001C-5494
: Sequence 5494, Application US/09134001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
: TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GTC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: PRIOR FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 5494
: LENGTH: 226
: TYPE: PRT
: ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5494
```

Alignment Scores:

Pred. No.: 78.6 Length: 226
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 4 Gaps: 0

US-09-817-318-1 (1-780) x US-09-134-001C-5494 (1-226)

QY 542 AACTGCAATGTAATAGATTG 522
|||||

Db 195 ThrValAsnValAsnArgLeu 201

RESULT 32

US-08-484-905-65
; Sequence 65, Application US/08484905
; Patent No. 5976551

; GENERAL INFORMATION:

; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre

; TITLE OF INVENTION: An Altered Major Histocompatibility

; TITLE OF INVENTION: Complex(MHC) Determinant and Methods for Using the

; NUMBER OF SEQUENCES: 127

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W., Suite 700; CITY: Washington
; STATE: D.C.

; ZIP: 20005-3315

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk
; OPERATING SYSTEM: PC-DOS-/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/484,905

; FILING DATE: 07-JUNE-1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/792,473

; FILING DATE: 15-NOV-1991

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Potter, Jane E. R.

; REGISTRATION NUMBER: 33,332

; REFERENCE/DOCKET NUMBER: 03495.0106-03000

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-408-4400

; TELEFAX: 202-408-4400

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 350 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FEATURE:

; NAME/KEY: Region

; LOCATION: 326..327

; OTHER INFORMATION: /note= "Xaa is an unidentified

; OTHER INFORMATION: amino acid residue."

US-08-484-905-65

Alignment Scores:

Pred. No.: 74.2 Length: 350
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.92% Indels: 0

DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x US-08-484-905-65 (1-350)

QY 625 ACTCTGATTATCCGCCCC 605
|||||

Db 331 ThrSerAspLeuSerLeuPro 337

RESULT 33

US-08-484-905-67
; Sequence 67, Application US/08484905
; Patent No. 5976551

; GENERAL INFORMATION:

; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre

; TITLE OF INVENTION: An Altered Major Histocompatibility

; TITLE OF INVENTION: Complex(MHC) Determinant and Methods for Using the

; NUMBER OF SEQUENCES: 127

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W., Suite 700; CITY: Washington
; STATE: D.C.

; ZIP: 20005-3315

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk
; OPERATING SYSTEM: PC-DOS-/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/484,905

; FILING DATE: 07-JUNE-1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/792,473

; FILING DATE: 15-NOV-1991

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Potter, Jane E. R.

; REGISTRATION NUMBER: 33,332

; REFERENCE/DOCKET NUMBER: 03495.0106-03000

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-408-4400

; TELEFAX: 202-408-4400

; INFORMATION FOR SEQ ID NO: 67:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 350 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FEATURE:

; NAME/KEY: Region

; LOCATION: 326..327

; OTHER INFORMATION: /note= "Xaa is an unidentified

; OTHER INFORMATION: amino acid residue."

US-08-484-905-67

Alignment Scores:

Pred. No.: 74.2 Length: 350
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.92% Indels: 0
DB: 2 Gaps: 0
US-09-817-318-1 (1-780) x US-08-484-905-67 (1-350)
QY 625 ACTTGTGATTATACCGCC 605
|||||
Db 331 ThrsrAspLeuSerLeuPro 337
RESULT 34
US-08-481-985B-65
; Sequence 65, Application US/08481985B
; Patent No. 6011146
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,985B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0106-04000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 350 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Region
; LOCATION: 326..327
; OTHER INFORMATION: /note="Xaa is an unidentified
; OTHER INFORMATION: amino acid residue."
US-08-481-985B-65
Alignment Scores:
Pred. No.: 74.2 Length: 350
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 3 Gaps: 0
US-09-817-318-1 (1-780) x US-08-481-985B-65 (1-350)

QY 625 ACTTGTGATTATACCGCC 605
|||||
Db 331 ThrsrAspLeuSerLeuPro 337
RESULT 35
US-08-481-985B-67
; Sequence 67, Application US/08481985B
; Patent No. 6011146
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,985B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0106-04000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 350 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Region
; LOCATION: 326..327
; OTHER INFORMATION: /note="Xaa is an unidentified
; OTHER INFORMATION: amino acid residue."
US-08-481-985B-67
Alignment Scores:
Pred. No.: 74.2 Length: 350
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 3 Gaps: 0
US-09-817-318-1 (1-780) x US-08-481-985B-67 (1-350)
QY 625 ACTTGTGATTATACCGCC 605
|||||
Db 331 ThrsrAspLeuSerLeuPro 337

RESULT 36
US-08-370-476-65
Sequence 65, Application US/08370476
Patent No. 6153408
GENERAL INFORMATION:
APPLICANT: Mottez, Estelle
APPLICANT: Abastado, Jean-Pierre
APPLICANT: Kourilsky, Philippe
APPLICANT: Lone, Yu-Chun
APPLICANT: Ojcius, David
APPLICANT: Casrouge, Armand
TITLE OF INVENTION: Altered Major Histocompatibility Complex
NUMBER OF SEQUENCES: 127
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
Dunnet
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370,476
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/117,575
FILING DATE: 07-SEP-1993
APPLICATION NUMBER: US 08/072,787
FILING DATE: 06-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/801,818
FILING DATE: 05-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/792,473
FILING DATE: 15-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05243.0001-01000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Region
LOCATION: 326..327
OTHER INFORMATION: /note="Xaa is an unidentified
OTHER INFORMATION: amino acid residue."
US-08-370-476-65
Alignment Scores:
Pred. No.: 74.2 Length: 350
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 4 Gaps: 0
US-09-817-318-1 (1-780) x US-08-370-476-65 (1-350)
QY 625 ACTTCGATTATCCTGCC 605

Db 331 ThrSerAspLeuSerLeuPro 337
|||||
RESULT 37
US-08-370-476-67
Sequence 67, Application US/08370476
Patent No. 6153408
GENERAL INFORMATION:
APPLICANT: Mottez, Estelle
APPLICANT: Abastado, Jean-Pierre
APPLICANT: Kourilsky, Philippe
APPLICANT: Lone, Yu-Chun
APPLICANT: Ojcius, David
APPLICANT: Casrouge, Armand
TITLE OF INVENTION: Altered Major Histocompatibility Complex
NUMBER OF SEQUENCES: 127
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
Dunnet
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370,476
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/117,575
FILING DATE: 07-SEP-1993
APPLICATION NUMBER: US 08/072,787
FILING DATE: 06-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/801,818
FILING DATE: 05-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/792,473
FILING DATE: 15-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05243.0001-01000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Region
LOCATION: 326..327
OTHER INFORMATION: /note="Xaa is an unidentified
OTHER INFORMATION: amino acid residue."
US-08-370-476-67
Alignment Scores:
Pred. No.: 74.2 Length: 350
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 4 Gaps: 0
US-09-817-318-1 (1-780) x US-08-370-476-67 (1-350)

OY 625 ACTTGATTATCCCTGCC 605
|||||
DB 331 ThreSerAspLeuSerLeuPro 337

RESULT 38

US-08-153-848-28
; Sequence 28, Application US/08153848
; Patent No. 5759804

GENERAL INFORMATION:

APPLICANT: Godiska, Ronald
APPLICANT: Gray, Patrick W.
APPLICANT: Schweikart, Vicki L.
TITLE OF INVENTION: No. 5759804e1 Seven Transmembrane Receptors
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Bicknell
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153.848
FILING DATE:
CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/977,452
FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5759804and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31794
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-153-848-28

Alignment Scores:

Pred. No.:	74.1	Length:	355
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.86%	Indels:	0
DB:	1	Gaps:	0

US-09-817-318-1 (1-780) x US-08-153-848-28 (1-355)

OY 704 TTAATCTCCGCTGTAATA 724
|||||

DB 232 LeuileLeuValValille 238

RESULT 39

US-08-153-848-32
; Sequence 32, Application US/08153848
; Patent No. 5759804

GENERAL INFORMATION:

APPLICANT: Godiska, Ronald
APPLICANT: Gray, Patrick W.
APPLICANT: Schweikart, Vicki L.

TITLE OF INVENTION: No. 5759804e1 Seven Transmembrane Receptors
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Bicknell
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153.848
FILING DATE:
CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/977,452
FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5759804and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31794
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-153-848-32

Alignment Scores:

Pred. No.:	74.1	Length:	355
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.86%	Indels:	0
DB:	1	Gaps:	0

US-09-817-318-1 (1-780) x US-08-153-848-32 (1-355)

OY 704 TTAATCTCCGCTGTAATA 724
|||||

DB 232 LeuileLeuValValille 238

RESULT 40

US-09-299-843A-28
; Sequence 28, Application US/09299843A
; Patent No. 6107475

GENERAL INFORMATION:

APPLICANT: Godiska, Ronald
APPLICANT: Gray, Patrick W.
APPLICANT: Schweikart, Vicki L.
TITLE OF INVENTION: No. 6107475e1 Seven Transmembrane Receptors
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,843A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/088,337
FILING DATE: 01-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/153,848
FILING DATE: 17-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,452
FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: JILL E. UHL
REGISTRATION NUMBER: 43,213
REFERENCE/DOCKET NUMBER: 27866/32059B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX:
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-299-843A-28

Alignment Scores:
Pred. No.: 74.1 Length: 355
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
Gaps: 0
DB: 3

US-09-817-318-1 (1-780) x US-09-299-843A-28 (1-355)

QY 704 TTAATTCCTCGTTGTAATA 724
|||||
DB 232 LeuileuValValIle 238

RESULT 41
US-09-299-843A-32
Sequence 32, Application US/09299843A
Patent No. 6107475
GENERAL INFORMATION:
APPLICANT: Godiska, Ronald
APPLICANT: Gray, Patrick W.
APPLICANT: Schweiakart, Vicki L.
TITLE OF INVENTION: No. 6107475el Seven Transmembrane Receptors
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,843A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/088,337
FILING DATE: 01-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/153,848
FILING DATE: 17-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,452
FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: JILL E. UHL
REGISTRATION NUMBER: 43,213
REFERENCE/DOCKET NUMBER: 27866/32059B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX:
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-299-843A-32

Alignment Scores:
Pred. No.: 74.1 Length: 355
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
Gaps: 0
DB: 3

US-09-817-318-1 (1-780) x US-09-299-843A-32 (1-355)

QY 704 TTAATTCCTCGTTGTAATA 724
|||||
DB 232 LeuileuValValIle 238

RESULT 42
US-09-088-337B-28
Sequence 28, Application US/09088337B
Patent No. 6348574
GENERAL INFORMATION:
APPLICANT: Godiska, Ronald
APPLICANT: Gray, Patrick W.
APPLICANT: Schweiakart, Vicki L.
TITLE OF INVENTION: No. 6348574el Seven Transmembrane Receptors
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/088,337B
FILING DATE: 01-Jun-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/153,848
FILING DATE: 17-NOV-1993
APPLICATION NUMBER: US 07/977,452
FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6348574and, Greta E.
REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER: 31794
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-817-318-1 (1-780) x US-09-088-337B-28 (1-355)
Alignment Scores:
Pred. No.: 74.1 Length: 355
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
Gaps: 0
DB: 4
US-09-817-318-1 (1-780) x US-09-088-337B-28 (1-355)
QY 704 TTAATCTCCTGCTTGAATA 724
Db 232 LeuileLeuValIle 238
RESULT 43
US-09-088-337B-32
Sequence 32, Application US/09088337B
Patent No. 6348574
GENERAL INFORMATION:
APPLICANT: Godiska, Ronald
Gray, Patrick W.
Schweikart, Vicki L.
TITLE OF INVENTION: No. 6348574el Seven Transmembrane Receptors
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/088,337B
FILING DATE: 01-Jun-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/153,848
FILING DATE: 17-NOV-1993
APPLICATION NUMBER: US 07/977,452
FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6348574and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31794
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-0448
TELEFAX: (312) 474-6300
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-088-337B-32
Alignment Scores:
Pred. No.: 74.1 Length: 355
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
Gaps: 0
DB: 4
US-09-817-318-1 (1-780) x US-09-088-337B-32 (1-355)
QY 704 TTAATCTCCTGCTTGAATA 724
Db 232 LeuileLeuValIle 238
RESULT 44
PCT-US93-11153-28
Sequence 28, Application PC/TUS9311153
GENERAL INFORMATION:
APPLICANT: Godiska, Ronald
Gray, Patrick W.
Schweikart, Vicki L.
TITLE OF INVENTION: Novel Seven Transmembrane Receptors
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
Bicknell
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/11153
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,452
FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Noland, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31794
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-11153-28
Alignment Scores:
Pred. No.: 74.1 Length: 355
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
Gaps: 0
DB: 5
US-09-817-318-1 (1-780) x PCT-US93-11153-28 (1-355)

QY	704	TTAATTCTCCTGGTTGTAATA	724
Db	232	LeuIleLeuLeuValValIle	238

RESULT 45

```

PCT-US93-11153-32
: Sequence 32, Application PC/TUS9311153
: GENERAL INFORMATION:
: APPLICANT: Godiska, Ronald
: APPLICANT: Gray, Patrick W.
: APPLICANT: Schweikart, Vicki L.
: TITLE OF INVENTION: Novel Seven Transmembrane Receptors
: NUMBER OF SEQUENCES: 64
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
: ADDRESSEE: Ricknell
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US93/11153
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/977,452
: FILING DATE: 17-NOV-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Noland, Greta E.
: REGISTRATION NUMBER: 35,302
: REFERENCE/DOCKET NUMBER: 31794
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 474-6300
: TELEFAX: (312) 474-0448
: TELEX: 25-3836
: INFORMATION FOR SEQ ID NO: 32:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 355 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
PCT-US93-11153-32

```

Alignment Scores:	
Pred. No.:	74.1
Length:	355
Score:	7.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	2.86%
DB:	5
Matches:	7
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-817-318-1 (1-780) x PCT-US93-11153-32 (1-355)

QY	704	TTAATTCCTCGTGTGTAATA	724
Db	232	LeuIleLeuLeuValIle	238

Search completed: February 4, 2003, 07:24:18
Job time : 28.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 4, 2003, 07:17:28 ; Search time 57.5 Seconds

(without alignments)
5590.148 Million cell updates/sec

Title: US-09-817-318-1

Perfect score: 245

Sequence: 1 aattactgtctctcttaag.....taagggcaaaaaaaaaaaaaa 780

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1343044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-O=/c9n2.1/USPTO_spool/US09817318/runat_04022003_071722_6623/app_query.fasta.1.967
-DB=SPTREMBL_21 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=quality -THR_MIN=1 -ALIGN=45 -MODE=LOCAL -OUTFMT=pio
-NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09817318.ecgn.1.1.88.@runat_04022003_071722_6623 -NCPU=6 -ICPU=3
-NO_XLPPY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=7
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

1: SPTREMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
C 1	11	4.6	350 5 Q8SV97	Q8SV97 encephalito

2	8	3.3	145	16	Q9PM11	Q9PM11 campylobact
3	8	3.3	168	17	Q8ZMK1	Q8ZMK1 pyrobaculum
C 4	8	3.3	235	10	Q65527	Q65527 arabidopsis
C 5	8	3.3	266	16	Q9RYR5	Q9RYR5 deinococcus
C 6	8	3.3	300	5	Q9NH08	Q9NH08 agrotis ips
C 7	8	3.3	311	17	Q8TX17	Q8TX17 methanopyru
9	8	3.3	341	10	Q957G2	Q957G2 arabidopsis
10	8	3.3	454	10	Q9ZT91	Q9ZT91 arabidopsis
C 11	8	3.3	866	16	Q8UCG6	Q8UCG6 agrobacteri
C 12	8	3.3	963	10	Q9M147	Q9M147 arabidopsis
C 13	8	3.3	1613	16	Q9KRZ1	Q9KRZ1 vibrio chol
C 14	7	2.9	21	8	Q35556	Q35556 pythium oli
C 15	7	2.9	70	12	Q991K5	Q991K5 hepatitis b
C 16	7	2.9	76	17	Q58627	Q58627 methanococ
C 17	7	2.9	90	8	Q9G425	Q9G425 acipenser g
C 18	7	2.9	91	9	Q48470	Q48470 bacterioph
C 19	7	2.9	100	8	Q78856	Q78856 phytophthor
C 20	7	2.9	102	7	Q31193	Q31193 mus musculi
C 21	7	2.9	139	8	Q32919	Q32919 phacelocar
C 22	7	2.9	139	4	Q96DR6	Q96DR6 homo sapien
C 23	7	2.9	146	9	Q34053	Q34053 streptococ
C 24	7	2.9	147	5	Q9XXA6	Q9XXA6 caenorhabdi
C 25	7	2.9	152	8	Q48024	Q48024 candida asp
C 26	7	2.9	152	12	Q9D0M9	Q9D0M9 kaposi's sa
C 27	7	2.9	153	16	Q9A3V4	Q9A3V4 caulobacter
C 28	7	2.9	164	7	Q31156	Q31156 mus musculi
C 29	7	2.9	168	16	Q99R41	Q99R41 staphylococ
C 30	7	2.9	170	8	Q9MR45	Q9MR45 echinococu
C 31	7	2.9	175	10	Q9LS47	Q9LS47 arabidopsis
C 32	7	2.9	175	16	Q9RX34	Q9RX34 delinococcus
C 33	7	2.9	188	5	Q9U0N2	Q9U0N2 plasmodium
C 34	7	2.9	189	5	Q9VYN0	Q9VYN0 drosophila
C 35	7	2.9	193	8	Q8WCU5	Q8WCU5 rhadophus
C 36	7	2.9	196	8	Q36064	Q36064 trichiturus
C 37	7	2.9	197	16	Q8B9K7	Q8B9K7 thelmoanaer
C 38	7	2.9	207	3	Q9HF54	Q9HF54 ashbya goss
C 39	7	2.9	211	4	Q9UFW5	Q9UFW5 homo sapien
C 40	7	2.9	219	10	Q9M2H9	Q9M2H9 arabidopsis
C 41	7	2.9	220	15	Q9WKZ4	Q9WKZ4 human immun
C 42	7	2.9	220	15	Q9WKZ3	Q9WKZ3 human immun
C 43	7	2.9	220	15	Q9WKX1	Q9WKX1 human immun
C 44	7	2.9	220	15	Q9KMW7	Q9KMW7 human immun
C 45	7	2.9	222	16	Q8XFN8	Q8XFN8 salmonella
			224	16	Q99VW2	Q99VW2 staphylococ

ALIGNMENTS

RESULT 1
ID Q8SV97 PRELIMINARY; PRT; 350 AA.
AC Q8SV97;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Hypothetical protein EC006_1070.
GN EC006_1070.
OS Encephalitozoon cuniculi.
OC Eukaryota; Microsporidia; Unikaryoniidae; Encephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA Genoscope;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RX MEDLINE=21576510; PubMed=11719806;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
RA Prensier G., Barbe V., Peyreallade E., Brotlier P., Wincker P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivares C.P.;

"Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi";
RL Nature 414:450-453(2001).
DR EMBL: AL590446; CAD25467.1; -
KW Hypothetical protein. 39829 MW; 239CBF411086C3BA CRC64;
SQ SEQUENCE 350 AA; 39829 MW; 239CBF411086C3BA CRC64;

Alignment Scores:
Pred. No.: 0.0276 Length: 350
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.58% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x Q8SV97 (1-350)

QY 336 CTGTCAAAATTCATGAGTCTTGACATTTCT 304
Db 134 LeuSerLysIleHISGLuValLeuAspTyrSer 144

RESULT 2

Q9PM11 PRELIMINARY; PRT; 145 AA.
AC Q9PM11:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Putative periplasmic protein.
GN CJI666C.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=NCYC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Kettle J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
RA Jaisels K., Kariyasek A.V., Moule S., Palen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences";
RL Nature 403:665-668(2000).
DR EMBL: AL139079; CAB73653.1; -
KW Complete proteome.
SQ SEQUENCE 145 AA; 16282 MW; 6BE64F3D3BC69BA CRC64;

Alignment Scores:
Pred. No.: 41.1 Length: 145
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x Q9PM11 (1-145)

QY 651 TATGGATTTTATAGTGATATAA 674
Db 136 TyrGlyTyrPheIleGlyAspLys 143

RESULT 3

Q8ZWK1 PRELIMINARY; PRT; 168 AA.
AC Q8ZWK1:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE P. aerophilum family 550 protein.
GN PAE1755.

OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL: AE009837; AAL63701.1; -
KW Complete proteome.

US-09-817-318-1 (1-780) x Q8ZWK1 (1-168)

Alignment Scores:
Pred. No.: 40.1 Length: 168
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x Q8ZWK1 (1-168)

QY 698 CATGTTTAACTCCGCTGTA 721
Db 111 HisValIleIleuValVal 118

RESULT 4

Q65527 PRELIMINARY; PRT; 235 AA.
AC Q65527:
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE Hypothetical 25.9 kDa protein.
GN F4D1.60 OR ATAG32740.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eudicots II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]

RP SEQUENCE FROM N.A.
RA Bevan M., Benes V., Rechmann S., Borkova D., Ansoorge W., Hohnel J.,
RA Mewes H.W., Mayer K.F.X., Schueller C.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Benes V., Rechmann S., Borkova D., Ansoorge W., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]

RP SEQUENCE FROM N.A.
RA Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL022537; CAA18587.1; -
DR EMBL: AL161582; CAB79991.1; -
KW Hypothetical protein.
SQ SEQUENCE 235 AA; 25859 MW; CBB9FF0207D292F8 CRC64;

Alignment Scores:
Pred. No.: 38 Length: 235
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x Q65527 (1-235)

```
OY 482 CATAAATCAGTTGGTGATTC 459
Db 21 H1SLYSLIESErLeuGlyAspPhe 28

RESULT 5
O9RYYS PRELIMINARY: PRT: 266 AA.
AC O9RYYS:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Molybdenum ABC transporter, permealase protein, putative.
GN DRA0168.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1:
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Halt D.H., Gwinn M.B., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Uitterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577(1999).
DR EMBL: AE001862; AAF12201.1; -.
DR TIGR: DRA0168; -.
DR InterPro: IPR00515; BPD_transp.
DR Pfam: PF00528; BPD_transp; 1.
KW Complete proteome.
SQ SEQUENCE 266 AA; 27799 MW; 673D6AEDADCE9548 CRC64;

Alignment Scores:
Pred. No.: 37.2 Length: 266
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
Gaps: 0

US-09-817-318-1 (1-780) x O9RYYS (1-266)
OY 646 ATAACTTGTTGACAGTACT 623
Db 61 lIethrleuLeuThrlValthr 68

RESULT 6
O9NH08 PRELIMINARY: PRT: 300 AA.
AC O9NH08:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE A1C6 chymotrypsinogen.
OS Agrotis ipsilon (Black cutworm moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Noctuidae; Noctuinae; Agrotis.
OX NCBI_TaxID=56364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MIDGUT;
RX MEDLINE=21167524; PubMed=11267902;
RA Mazumdar-Leighton S., Broadway R.M.;
RT "Identification of six chymotrypsin cDNAs from larval midguts of
RT Helicoverpa zea and Agrotis ipsilon feeding on the soybean (Kunitz)
```

```
RT trypsin inhibitor."
RL Insect Biochem. Mol. Biol. 31:633-644(2001).
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
DR EMBL: AF233730; AAF71517.1; -.
DR HSSP: P00763; IDPO.
DR MEROPS: S01.121; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease TRY.
DR Pfam: PF00089; Trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPsin.
DR SMART: SM00020; Tryp_spec; 1.
DR PROSITE: PS50240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
KW Hydrolyase; Serine protease.
FT CHAIN 61 300
SQ SEQUENCE 300 AA; 31418 MW; CEDBFAEA2FC7C6 CRC64;

Alignment Scores:
Pred. No.: 36.5 Length: 300
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
Gaps: 0

US-09-817-318-1 (1-780) x O9NH08 (1-300)
OY 323 ATGAGTCTTGACTATCTTTG 300
Db 1 Melysphleuthrilleleu 8

RESULT 7
O8TX17 PRELIMINARY: PRT: 311 AA.
AC O8TX17:
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Phosphate uptake regulator.
GN PHOU OR MK0862.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,
RA Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malikh A.G., Koonin E.V., Kozlovskiy S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR EMBL: AE010376; AAM02075.1; -.
KW Complete proteome.
SQ SEQUENCE 311 AA; 35909 MW; F3DD03066D43CD09 CRC64;

Alignment Scores:
Pred. No.: 36.3 Length: 311
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
Gaps: 0

US-09-817-318-1 (1-780) x O8TX17 (1-311)
OY 34 GGTGAAGCCTTACTTAAAGA 11
|||||
```

Db 88 GlyValArgProTyrPheLysArg 95

RESULT 8

Q9S7G2 PRELIMINARY: PRT: 341 AA.

AC Q9S7G2

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Hypothetical 38.8 kDa protein.

GN T3A5.140 OR F18B3.40.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Bloembergen H., Mewes H.W., Lemcke K., Mayer K.F.X., Quetler F.,

RA Salanoubat M.;

RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Rieger M., Gabel C., Mueller-Auer S., Schaefer M., Zipp M.,

RA Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C., Quetler F.,

RA Salanoubat M.;

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

RL EMBL: AL132979; CAB62445.1; -

DR EMBL: AL049862; CAB42905.1; -

DR InterPro: IPR002495; GI_8.

DR Pfam: PF01501; Glyco_Transf.8; 1.

KW Hypothetical protein; Transferase.

SEQUENCE 341 AA; 38774 MW; 96A527DFC5FB9E CRC64;

SO

Alignment Scores:

Pred. No.:	35.7	Length:	341
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.27%	Indels:	0
DB:	10	Gaps:	0

US-09-817-318-1 (1-780) x Q9S7G2 (1-341)

QY 98 AGTAGTTATTTATATATATTC 121

Db 3 SerIysPheLeuTyrLeuSer 10

RESULT 9

Q9ZT91 PRELIMINARY: PRT: 454 AA.

AC Q9ZT91

DT 01-MAY-1999 (TREMBlrel. 10, Created)

DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Elongation factor Tu, mitochondrial precursor.

GN TUF4 OR AT4G02930 OR T419.19 OR T5J8.25.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN-CV. COLUMBIA;

RX MEDLINE=96145515; PubMed=8555448;

RA Kuhlman P., Palmer J.D.;

RT "Isolation, expression, and evolution of the gene encoding

RT mitochondrial elongation factor Tu in Arabidopsis thaliana.";

RL Plant Mol. Biol. 23:1057-1070(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-CV. COLUMBIA;

RX MEDLINE=20083488; PubMed=10617198;

RA Mayer K.F.X., Schueller C., Mambut R., Murphy G., Volckaert G.,

RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.D., Teyrn N.,

RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,

RA Weichselgartner M., de Simone V., Obermair B., Maché R., Mueller M.,

RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,

RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,

RA Vos P., Hohenseil J., Zimmermann W., Weller H., Ridley P.,

RA Langham S.-A., McCullagh B., Bilham L., Robben J.,

RA Van der Schueren J., Grymonprez B., Chang Y.-J., Vandenbussche F.,

RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,

RA Weltzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,

RA Holzner E., Brandt A., Peters S., van Staveren M., Dirke W.,

RA Moeljan P., Klein Lankhorst R., Rose M., Hauf U., Koetter P.,

RA Berner S., Hempel S., Feldpausch M., Lambers S., Van den Daele H.,

RA de Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,

RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,

RA Clark L., Doggett J., Hall S., Kay M., Leonard N., McLeay K., Mayes R.,

RA Pettett A., Rajendram M.-A., Lyne M., Benes V., Reumann S.,

RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehert T.-H.,

RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs W., Fairman B., Grandtack K., Danner D., Hertzl A.,

RA Neumann A., Argitliou A., Vitale D., Liguori R., Piravandi E.,

RA Massenot O., Quigley F., Clapaud G., Muendlein A., Felber R.,

RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,

RA Cheifod F., Cooke R., Berger C., Monfort A., Cascuperta E.,

RA Gibbons T., Weber N., Vandenbol M., Baryes M., Terol J., Torres A.,

RA Perez-Perez A., Punnelle B., Bent E., Johnson S., Jacou D., Jesse T.,

RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielek C.,

RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,

RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,

RA Parnell L., Dedhia N., Gnoj L., Schütz K., Huang E., Spiegel L.,

RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,

RA Stoneking T., Kallick J., Graves T., Harmon G., Edwards J.,

RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,

RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,

RA Nelson J., Fulton L., Mardis E., Dante M., Geipel K., Hillier L.,

RA Krasner J., Spiech J., Ryan E., Andrews S., Giesel C., Layman D.,

RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,

RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,

RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,

RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,

RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,

RA Chen E., Marra M., Martienssen R., McCombie W.R.;

RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis

RT thaliana.";

RL Nature 402:769-777(1999).

RN [3]

RP SEQUENCE FROM N.A.

RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,

RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,

RA Shekher M., Schütz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,

RA Mewes H.W., Lemcke K., Mayer K.F.X.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE OF 342-454 FROM N.A.

RC STRAIN-CV. COLUMBIA;

RA Cooke R., Landie M., Raynal M., Delseny M.;

RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.

RN [5]

RP SEQUENCE OF 296-306, AND SUBCELLULAR LOCATION.

RC TISSUE-LEAF, AND STEM;

RX MEDLINE=21608403; PubMed=11743114;

RA Krut V., Eubel H., Jaensch L., Werhahn W., Braun H.-P.;

RT "Proteomic approach to identify novel mitochondrial proteins in


```

RT Arabidopsis.
RL Plant Physiol. 127:1694-1710(2001).
CC -i- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-tRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
DR EMBL: X89227; CAA61511.1; -.
DR EMBL: AF069442; AAC79113.1; -.
DR EMBL: AC004044; AAD15337.1; -.
DR EMBL: AL161495; CAB7778.1; -.
DR EMBL: F14375; CAA23078.1; -.
DR HSSP: P02990; 1EFU.
DR InterPro: IPR004541; EF-Tu.
DR InterPro: IPR004160; EFTU_Cterm.
DR InterPro: IPR004161; EFTU_D2.
DR InterPro: IPR000795; EF-GTPbind.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR PRINTS: PR00315; ELONGATNFACT.
DR TIGRFAMS: TIGR00485; EF-Tu; 1.
DR TIGRFAMS: TIGR00231; small_GTP; 1.
DR PROSITE: PS00301; EFACITOR_GTP; 1.
KW Elongation factor; GTP-binding; Protein biosynthesis; Transit peptide;
KW Mitochondrion.
FT TRANSIT 1 51 MITOCHONDRION (POTENTIAL).
FT CHAIN 52 454 ELONGATION FACTOR TU.
FT NP_BIND 74 81 GTP (BY SIMILARITY).
FT NP_BIND 136 140 GTP (BY SIMILARITY).
FT NP_BIND 191 194 GTP (BY SIMILARITY).
FT CONFLICT 1 10 MASVYLRLNPS -> MEEPVDEIDILRVSSTDIGMNEFAAA
FT (IN REF. 1).
FT CONFLICT 310 312 VPL -> CST (IN REF. 1).
SQ SEQUENCE 454 AA; 49409 MW; B7C2CAlDCC432FDA CRC64;

Alignment Scores:
Pred. No.: 34.1 Length: 454
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 10 Gaps: 0

US-09-817-318-1 (1-780) x Q9ZT91 (1-454)
QY 587 AAAAGATATTGATTAATGGCGAG 610
| | | | | | | | | | | | | | | | | |
Db 323 Lysylsileleuaspanglycin 330

RESULT 10
Q8UCG6 PRELIMINARY: PRT; 866 AA.
ID 08UCG6
AC 08UCG6;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical protein Atu2521.
GN Atu2521 OR AGR_C_4580.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria: Proteobacteria; alpha subdivision: Rhizobiales group;
OC Rhizobiales; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D., Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayvin T., Levy R., Li M.-J., McClelland E., Palmeri A.,

```

```

RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorllo B., Goldman B.S., Cao Y., Akenazi M., Halling C., Mullin L.,
RA Houmlel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Woliam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Chelo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
DR EMBL: AE009198; AAL43508.1; -.
DR EMBL: AE008164; AAK88252.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 866 AA; 93876 MW; ID21156AA010EPCE CRC64;

Alignment Scores:
Pred. No.: 30.7 Length: 866
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 16 Gaps: 0

US-09-817-318-1 (1-780) x Q8UCG6 (1-866)
QY 428 GTTCCTTGCTATTAATGAACA 451
| | | | | | | | | | | | | | | | | |
Db 376 ValProleuproilleleluA1a 383

RESULT 11
Q9M147 PRELIMINARY: PRT; 963 AA.
ID 09M147
AC 09M147;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Hypothetical 111.4 kDa protein.
GN At4G01210.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL161491; CAB80930.1; -.
DR InterPro: IPR001296; Glycos_Transf_1.
DR Pfam: PF00534; Glycos_Transf_1; 1.
KW Hypothetical protein.
SQ SEQUENCE 963 AA; 111437 MW; DB865E093651B184 CRC64;

Alignment Scores:
Pred. No.: 30.1 Length: 963
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

```

Query Match: 3.33% Indels: 0
DB: 10 Gaps: 0

US-09-817-318-1 (1-780) x Q9M147 (1-963)

QY 348 AGAAGAGTCTGCTCAAAATT 325
|||||
DB 730 Arglysglusterleuserlysile 737

RESULT 12
Q9KRZ1 PRELIMINARY: PRT: 1613 AA.
AC Q9KRZ1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
GN VC1492.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
Ermolaeva M.D., Vamathevan J., Bass S., Olin H., Dragoi I., Sellers P.,
McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae.";
RL Nature 406:477-483(2000).
DR EMBL: AE004227; AAF94647.1;
DR TIGR: VC1492;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1613 AA; 183558 MW; 58902452A44A6583 CRC64;

Alignment Scores:
Pred. No.: 27.7 Length: 1613
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 16 Gaps: 0

US-09-817-318-1 (1-780) x Q9KRZ1 (1-1613)

QY 254 AGAGTAAGTCTGTAAAGCTGTCA 231
|||||
DB 264 Argvalargservalylser 271

RESULT 13
Q35556 PRELIMINARY: PRT: 21 AA.
ID Q35556;
AC Q35556;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 2.5 kDa protein.
OS Pythium oligandrum.
OC Mitochondrion.
OC Eukaryota; stramenopiles; Competes; Pythiales; Pythiaceae; Pythium.
OX NCBI_TaxID=41045;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17-1;
RA Martin F.N.;
RT "Linear mitochondrial genome organization in vivo in the genus
Pythium.";
RL Curr. Genet. 27:0-0(1995).

DR EMBL: U28355; AAA70034.1; -
KW Hypothetical protein; Mitochondrion.
SQ SEQUENCE 21 AA; 2491 MW; 883B7AB7A5D4445E CRC64;

Alignment Scores:
Pred. No.: 613 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 8 Gaps: 0

US-09-817-318-1 (1-780) x Q35556 (1-21)

QY 114 TATAAATAACTTACTGTCA 94
|||||
DB 14 Tyrlsileasleuser 20

RESULT 14
Q991K5 PRELIMINARY: PRT: 70 AA.
ID Q991K5;
AC Q991K5;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Truncated precore/core protein.
OS Hepatitis B virus.
OC Viruses; Retrod viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RA Dong J., Cheng J., Wang Q., Shi S., Zhang J., Xia X., Si C.;
RT "HBV quasiespecies: prec/C region as example.";
RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF335737; AKK30081.1; -
DR InterPro: IPR002006; Hepatitis_core.
DR Pfam: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 70 AA; 7826 MW; 1460F931A854ABB CRC64;

Alignment Scores:
Pred. No.: 504 Length: 70
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 12 Gaps: 0

US-09-817-318-1 (1-780) x Q991K5 (1-70)

QY 404 TTGTTACTAGTTTCTCTG 424
|||||
DB 49 Leuleuthrserpheuleu 55

RESULT 15
Q58627 PRELIMINARY: PRT: 76 AA.
ID Q58627;
AC Q58627;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein MJ1230.
GN MJ1230.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanococcaceae; Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,

RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geophagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, *Methanococcus*
 RT *jannaschii*.";
 RL Science 273:1058-1073(1996).
 DR EMBL: U67563; AAB99236.1; -.
 DR TIGR: MJ1230; -.
 DR InterPro: IPR001052; Rubredoxin.
 DR PROSITE: PS00202; RUBREDOXIN; UNKNOWN_1.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 76 AA; 8903 MW; 2C53E9EABEF0247C CRC64;

Alignment Scores:

Pred. No.:	497	Length:	76
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	17	Gaps:	0

US-09-817-318-1 (1-780) x Q58627 (1-76)

QY 267 AAGTCAATTGTGAAGTAG 247

DB 58 LysSerIleValLysSerLys 64

RESULT 16

ID 09G425 PRELIMINARY: PRT: 90 AA.

AC 09G425; 01-MAR-2001 (TReMBLrel. 16, Created)

DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

DE Cytochrome b (Fragment).

GN CYTB.

OS Acipenser guldenstadti (Caspian sturgeon) (Russian sturgeon).

OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;

OC Acipenser.

OX NCBI_TaxID=7902;

RN [1]

RC SEQUENCE FROM N.A.

RA STRAIN=8519L2B024; TISSUE=CAVIAR;

RA Fain S.R., Lemay J.P., Shafer J.A., Hoesch R.M., Hamlin B.C.,

RA Straughan D.J.;

RT "DNA Sequence Identification of Sturgeon Caviars Traveling in World

RT Trade.";

RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: COMPONENT OF THE UBIQUITIN-CTOCHROME C REDUCTASE

CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A

CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL

CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).

CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY

CC BOUND TO THE PROTEIN (BY SIMILARITY).

CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,

CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).

CC EMBL: AF308902; AAC47741.1; -.

DR InterPro: IPR000179; Cyt_b.b6.

DR Pfam: PF00033; Cytochrome_b_N.1.

DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.

KW Electron transport; Heme; Mitochondrion; Respiratory chain;

KW Transmembrane.

RT NON_TER 1

FT NON_TER 90

SQ SEQUENCE 90 AA; 10228 MW; 78166B021859BCA7 CRC64;

Alignment Scores: 483 Length: 90

Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	8	Gaps:	0

US-09-817-318-1 (1-780) x Q9G425 (1-90)

QY 780 TTTTGTGCTTGTAC 760

DB 54 PhePhePheCysLeuTy 60

RESULT 17

ID 048470 PRELIMINARY: PRT: 91 AA.

AC 048470; 01-JUN-1998 (TReMBLrel. 06, Created)

DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

DE Complete nucleotide sequence.

GN Bacteriophage SP1.

OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;

OC Lambda-like viruses.

OX NCBI_TaxID=10724;

RN [1]

RP SEQUENCE FROM N.A.

RA Alonso J.C., Juder G., Stege A.C., Chai S., Weise F., Trautner T.A.;

RT "Analysis of the complete nucleotide sequence and functional

RT organization of *Bacillus subtilis* bacteriophage SP1."

RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: X97918; CA65517.1; -.

SO SEQUENCE 91 AA; 11090 MW; B326ADE56F21505 CRC64;

Alignment Scores:

Pred. No.:	482	Length:	91
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	9	Gaps:	0

US-09-817-318-1 (1-780) x 048470 (1-91)

QY 342 GAGTCACTGCAAAATTCAT 322

DB 41 GluslerleuSerLysIleHis 47

RESULT 18

ID 078856 PRELIMINARY: PRT: 100 AA.

AC 078856; 01-NOV-1998 (TReMBLrel. 08, Created)

DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

DE NADH dehydrogenase subunit 4L.

GN NAD4L.

OS Phytophthora cinnamomi.

OG Mitochondrion.

OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;

OC Phytophthora.

OX NCBI_TaxID=4785;

RN [1]

RC SEQUENCE FROM N.A.

RA STRAIN=A2400;

RA MEDLINE=98336900; PubMed=9673029;

RA Dobrowolski M.P., Tommerup I.C., O'Brien P.A.;

RT "Mitochondrial genome of phytophthora cinnamomi

RT failed to provide highly polymorphic markers for population

RT genetics.";

RT FEMS Microbiol. Lett. 163:243-248(1998).

DR EMBL: AF051359; AAC39493.1; -.

DR InterPro: IPR003215; NADH_dh_ubiq1.

DR InterPro: IPR001133; Oxidored_4L.

DR InterPro: IPR003214; Oxidred4L.
 DR Pfam: PF00420; Oxidored_42; 1.
 DR ProDom: PD000359; Oxidred4L; 1.
 DR ProDom: PD002107; NADH_ch_ubidq; 1.
 KW Mitochondrion.
 SQ SEQUENCE 100 AA; 11311 MW; 8909925088B4C4E5 CRC64;

Alignment Scores:

Pred. No.: 475 Length: 100
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.92% Indels: 0
 DB: 8 Gaps: 0

US-09-817-318-1 (1-780) x O78856 (1-100)

OY 114 TATAAATTAACCTTACTCTCA 94
 DB 93 TylsileasleuSer 99
 |||||

RESULT 19

O31193 PRELIMINARY; PRT; 102 AA.

AC O31193; 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE MHC class I H-2K gene (Q haplotype), clone p13 (Exons 4-8) mRNA
 GN H2-K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84068207; PubMed=6689056;
 RA Kress M., Glaros D., Khoury G., Jay G.;
 RT "Alternative RNA splicing in expression of the H-2K gene.";
 RL Nature 306:602-604(1983).
 DR EMBL: K01762; AAA39655.1; -.
 DR MGD: MGI:95904; H2-K.
 DR InterPro: IPR003006; IG_MHC.
 DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
 FT NON_TER
 SQ SEQUENCE 102 AA; 10797 MW; B0D124B792B150C5 CRC64;

Alignment Scores:

Pred. No.: 473 Length: 102
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.92% Indels: 0
 DB: 7 Gaps: 0

US-09-817-318-1 (1-780) x O31193 (1-102)

OY 625 ACTTCGATTATGCCCTGCC 605
 DB 92 ThrsrAspSerLeuPro 98
 |||||

RESULT 20

O32919 PRELIMINARY; PRT; 129 AA.

AC O32919; 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Ribulose biphosphate carboxylase large chain (EC 4.1.1.39) (Rubisco
 GN RBCU.
 OS Phaeolocarpus tortuosus.

OC Chloroplast.
 OC Eukaryota; Rhodophyta; Florideophyceae; Gigartinales;
 OC Phaeolocarpaceae; Phaeolocarpus.
 OX NCBI_TaxID=40405;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Frederick S., Homersand M.H., Freshwater D.W.;
 RT "Erratum: The molecular systematics of some agar- and carrageenan-
 containing marine red algae based on rbcL sequence analysis.";
 RL Hydrobiologia 341:251-251(1996).

RP SEQUENCE FROM N.A.

RA Frederick S., Homersand M.H., Freshwater D.W.;
 RT "The molecular systematics of some agar- and carrageenan-
 containing marine red algae based on rbcL sequence analysis.";
 RL Hydrobiologia 326:125-135(1996).

CC -1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF D-
 CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
 THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
 REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
 ACTIVE SITE (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) = 3-
 PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.

CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) = 2 3-
 PHOSPHO-D-GLYCERATE.

CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: CHLOROPLAST (BY SIMILARITY).

DR EMBL: U26826; AAC36421.1; -.
 DR InterPro: IPR000685; Rubisco_large.

DR Pfam: PF00016; Rubisco_large_N; 1.

KW Carbon dioxide fixation; Chloroplast; Lyase; Monooxygenase;
 FT Oxidoreductase; Photorespiration; Photosynthesis.

FT NON_TER
 SQ SEQUENCE 129 AA; 13964 MW; 71FC15D65269B9E CRC64;

Alignment Scores:

Pred. No.: 456 Length: 129
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.92% Indels: 0
 DB: 8 Gaps: 0

US-09-817-318-1 (1-780) x O32919 (1-129)

OY 453 ACTGCTTCATATAGGCANA 433
 DB 64 ThrsrAspSerLeuPro 70
 |||||

RESULT 21

O96DR6 PRELIMINARY; PRT; 139 AA.

AC O96DR6; 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE CSGEF.
 GN SGEF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PROSTATIC CARCINOMA;
 RA Qi H., Fillion C., Labrie Y., Grenier J., Fournier A., Labrie C.;
 RT "Isolation and androgen regulation of human CSGEF, a splicing variant
 of a new putative member (SGEF) of Dbl family", that maps to 3q25.31.";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBD databases.
 DR EMBL: AF415176; AAU27002.1; -.

DR InterPro: IPR001849: PH.
 DR InterPro: IPR001452: SH3.
 DR Pfam: PF00018: SH3; 1.
 DR ProDom: PD000066: SH3; 1.
 DR PROSITE: PSS0003: PH_DOMAIN; 1.
 DR PROSITE: PSS0002: SH3; 1.
 SO SEQUENCE 139 AA; 15655 MW; 17A9712161A57CC5 CRC64;

Alignment Scores:

Pred. No.: 450 length: 139
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best local Similarity: 100.00% Mismatches: 0
 Query Match: 2.86% Indels: 0
 DB: Gaps: 0

US-09-817-318-1 (1-780) x Q96DR6 (1-139)

OY 526 CTATTACATGACTGACTT 546

Db 12 LeuphethrLeuphValleu 18

RESULT 22

O34053 PRELIMINARY; PRT; 146 AA.
 AC 034053;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ORF22.
 OS Streptococcus thermophilus temperate bacteriophage O1205.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
 OX NCBI_TaxID=5154;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CNR21205;
 RX MEDLINE=98048466; Pubmed=9387220;
 RA Stanley E., Fitzgerald G.F., Le Marrec C., Fayard B., van Sinderen D.;
 RT "sequence analysis and characterization of phi O1205, a temperate
 bacteriophage infecting Streptococcus thermophilus CNR21205.";
 RL Microbiology 143:0-0(0).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CNR21205;
 RA Stanley E., Fitzgerald G.F., Le Marrec C., Fayard B., van Sinderen D.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U88974; AAC79538.1; -;
 SO SEQUENCE 146 AA; 16161 MW; CB36026042F8D5BB CRC64;

Alignment Scores:

Pred. No.: 447 length: 146
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best local Similarity: 100.00% Mismatches: 0
 Query Match: 2.86% Indels: 0
 DB: Gaps: 0

US-09-817-318-1 (1-780) x O34053 (1-146)

OY 295 GTTTACAAAGATGTCAG 315

Db 34 ValTylrlysaTgIleVallys 40

RESULT 23

O9XXA6 PRELIMINARY; PRT; 147 AA.
 AC 09XXA6;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE 6R55.2 protein.
 GN 6R55.2.
 OS Caenorhabditis elegans.

OC Eukaryota: Metazoa: Nematoda: Chromadorea: Rhabditida: Rhabditoidea;
 OC Rhabditidae: Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wallis J.M.;
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; Pubmed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL: AL031222; CAA20210.1; -;
 SO SEQUENCE 147 AA; 16637 MW; 2EEDC828C57F95C4 CRC64;

Alignment Scores:

Pred. No.: 446 length: 147
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best local Similarity: 100.00% Mismatches: 0
 Query Match: 2.86% Indels: 0
 DB: Gaps: 0

US-09-817-318-1 (1-780) x O9XXA6 (1-147)

OY 415 TTTTCTCTGAGTCTCTTT 435

Db 114 PheSerValSerSerPhe 120

RESULT 24

O48024 PRELIMINARY; PRT; 152 AA.
 AC 048024;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Cytochrome b (Fragment).
 GN CYTB.
 OS Candoia aspera.
 OS Mitochondrion.
 OC Eukaryota: Metazoa: Chordata: Graniata: Vertebrata: Euteleostomi;
 OC Lepidosauria: Squamata: Scleroglossa: Serpentes; Boidea; Boidae;
 OC Candoia.
 OX NCBI_TaxID=51853;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Campbell B.N.;
 RT "Hic Sunt Serpentes - Molecular phylogenetics and the Boide
 RT (Serpentes: Boidea).";
 RL Thesis (1997), Queen's University, Kingston, Ontario, Canada.
 CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B366) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN (BY SIMILARITY).
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CC CYTOCHROME C1 AND THE RISKE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 DR EMBL: U69752; AAC01786.1; -;
 DR InterPro: IPR000179; Cyt_b6.
 DR Pfam: PF00033; cytochrome_b_n; 1.
 KW Electron transport; Heme; Mitochondrion; Respiratory chain;
 FT Transmembrane.
 FT NON_TER 1 1
 FT NON_TER 152 152
 SO SEQUENCE 152 AA; 17200 MW; BF19D28F22A49DCB CRC64;

Alignment Scores:

Pred. No.: 444 length: 152
 Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.92% Indels: 0
 DB: 8 Gaps: 0

US-09-817-318-1 (1-780) x 048024 (1-152)

QY 646 ATAACTTGTATTGACAGTA 626
 DB 129 IletHrLeuLeuThrVal 135

RESULT 25

ID Q9DUM9 PRELIMINARY; PRT; 152 AA.

AC Q9DUM9; PRELIMINARY; PRT; 152 AA.
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DE ORF K10.6.
 OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 OX NCBI_TaxID=37296;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20392178; PubMed=10933732;
 RA Lubyova B., Pitcha P.M.;
 RT "Characterization of a novel human herpesvirus 8-encoded protein,
 RT VIRF-3, that shows homology to viral and cellular interferon
 RT regulatory factors";
 RL J. Virol. 74:8194-8201(2000).
 DR EMBL: AF254765; AAF98389.1; -;
 SQ SEQUENCE 152 AA; 17801 MM; 9492017B2BCA7076 CRC64;

Alignment Scores:

Pred. No.: 444 Length: 152
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.86% Indels: 0
 DB: 12 Gaps: 0

US-09-817-318-1 (1-780) x Q9DUM9 (1-152)

QY 492 GCTTTGGAAGATTGCAC 512
 DB 77 AlaPhelylystleCystThr 83

RESULT 26

ID Q9A3V4 PRELIMINARY; PRT; 153 AA.

AC Q9A3V4; PRELIMINARY; PRT; 153 AA.
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DE Hypothetical protein CC3097.
 GN CC3097.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 19089 / CB15;
 MEDLINE=21173698; PubMed=11259647;
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Sherry J., Berry K.,
 RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL: AE005973; AAK25059.1; -;
 DR TIGR: CC3097; -;
 KW Hypothetical protein; Complete Proteome.
 SQ SEQUENCE 153 AA; 16852 MM; 7DD975EB39A62A4C CRC64;

Alignment Scores:

Pred. No.: 443 Length: 153
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.86% Indels: 0
 DB: 16 Gaps: 0

US-09-817-318-1 (1-780) x Q9A3V4 (1-153)

QY 227 TGTGTGACGCTTACAGACC 247
 DB 144 CysValThrAlaLeuGlnThr 150

RESULT 27

ID Q31156 PRELIMINARY; PRT; 164 AA.

AC Q31156; PRELIMINARY; PRT; 164 AA.
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE MHC class I H2 antigen gene (Haplotype d) (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=82150234; PubMed=6278432;
 RA Lalanne J.-L., Breggère F., Delarbre C., Abastado J.-P., Gachelin G.,
 RA Kourilsky P.;
 RT "Comparison of nucleotide sequences of mRNAs belonging to the mouse h-
 RT 2 multigene family";
 RL Nucleic Acids Res. 10:1039-1049(1982).
 DR EMBL: J00395; AAA39579.1; -;
 DR InterPro: IPR003587; Ig_C1.
 DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00047; Ig_1.
 DR SMART: SM00407; IgC1; 1.
 DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 164 AA; 17774 MM; 354250B6C9B3D066 CRC64;

Alignment Scores:

Pred. No.: 438 Length: 164
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.92% Indels: 0
 DB: 7 Gaps: 0

US-09-817-318-1 (1-780) x Q31156 (1-164)

QY 625 ACTTGTGATTATCCCTGCC 605
 DB 145 ThrSerAspLeuSerLeuPro 151

RESULT 28

ID Q99RA1 PRELIMINARY; PRT; 168 AA.

AC Q99RA1; PRELIMINARY; PRT; 168 AA.
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DE Truncated-SA protein.
 GN TRUNCATED-SA OR SA2389.
 OS Staphylococcus aureus (strain N315).

```
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus;
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Ito T., Kanamori M.,
RA Matsunaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yanashita A., Oshima K., Furuya K., Yoshino C., Shibata T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003137; BAB43694.1; -.
KW Complete proteome.
SQ SEQUENCE 168 AA; 18967 MW; D6B388B687BDDB CRC64;

Alignment Scores:
Pred. No.: 436 Length: 168
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x Q9MR45 (1-168)
OY 704 TTATTCCTCGTGTATATA 724
Db 10 LeuLeuLeuValValIle 16

RESULT 29
Q9MR45 PRELIMINARY; PRT; 170 AA.
AC Q9MR45;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE ATPase subunit 6.
GN ATP6.
OS Echinococcus vogeli.
OC Mitochondrion.
OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
OC Cyclophyllidae; Taeniidae; Echinococcus.
OX NCBI_TaxID=6213;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=South America;
RA Fukunaga M., Yokoyama N., Ito A., Nakao M.;
RT "Distribution of Echinococcus multilocularis in Hokkaido, Japan and
RT its geographical migration."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Le T.H., Pearson M.S., Blair D., Dai N., Zhang L.H., McManus D.P.;
RT "Complete mitochondrial genomes confirm the distinctiveness of the
RT horse-dog and sheep-dog strains of Echinococcus granulosus."
RL Parasitology 0:0-0(2002).
DR EMBL: AB027560; BAA95343.1; -.
DR EMBL: AY056612; AAL55466.1; -.
KW Mitochondrion.
SQ SEQUENCE 170 AA; 19726 MW; 948C17E0ED8A6DD2 CRC64;

Alignment Scores:
Pred. No.: 436 Length: 170
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: Gaps: 0
```

```
US-09-817-318-1 (1-780) x Q9MR45 (1-170)
OY 415 TTTCTCTGTGAGTTCTTT 435
Db 72 PheSerSerValSerSerPhe 78

RESULT 30
Q9LS47 PRELIMINARY; PRT; 175 AA.
AC Q9LS47;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Calcineurin b subunit (protein phosphatase 2b regulatory subunit)-like
DE protein (Hypothetical 20.0 kDa protein).
GN At3g18430.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty pl and TAC
RT clones."
RL DNA Res. 7:131-135(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H.,
RA Hayashizaki Y., Ishida J., Jones T., Kamita A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.;
RA Theologis A.;
RT "Full length cDNA of gene At3g18430 (GI:15229640).";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamita A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinzaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB026658; BAB01109.1; -.
DR EMBL: AY063789; AAL36096.1; -.
DR EMBL: AY091287; AAM14226.1; -.
DR HSSP: P06705; IAU1.
DR InterPro: IPR002048; EF-hand.
DR Pfam: PF00036; EF-hand; 2.
DR ProDom: PD000012; EF-hand; 1.
DR ProSITE: PS00018; EF_HAND; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 175 AA; 20032 MW; B26D6F2ADA26AFED CRC64;

Alignment Scores:
Pred. No.: 433 Length: 175
```

Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 10 Gaps: 0

US-09-817-318-1 (1-780) x Q9L547 (1-175)

OY 335 AGTGAAGCTTTCTAACCTT 355
DB 148 Seraspserphelutheleu 154

RESULT 31

O9RX34 PRELIMINARY; PRT: 175 AA.

AC O9RX34: 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
GN DR0481.
OS Deinococcus radiodurans.

OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-RI:
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,

RA Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Uterback T., Zaleski C.,
RA Maratova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.,

RT "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans RI."

RL Science 286:1571-1577(1999).

DR EMBL: AE001907; AAF10064.1; -

DR TIGR: DR0481; -

KW Hypothetical protein; Complete proteome

SO SEQUENCE 175 AA; 19419 MW; 7B8276EABAD5A890 CRC64;

Alignment Scores:

Pred. No.: 433 Length: 175
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 16 Gaps: 0

US-09-817-318-1 (1-780) x Q9RX34 (1-175)

OY 701 GTTTAATTCCTCGTGTGTA 721

DB 4 ValLeuLeuLeuValVal 10

RESULT 32

O9U0N2 PRELIMINARY; PRT: 188 AA.

AC O9U0N2: 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
GN MALP3.05.
OS Plasmodium falciparum (isolate 3D7).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-3D7;

RA Bowman S., Churcher C., Harris B., Harris D., Lawson D., Quail M.,
RA Barrell B.,
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL031746; CAB63560.1; -

KW Hypothetical protein

SO SEQUENCE 188 AA; 21896 MW; 7E42C44E9D1C7321 CRC64;

Alignment Scores:

Pred. No.: 428 Length: 188
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 5 Gaps: 0

US-09-817-318-1 (1-780) x Q9U0N2 (1-188)

OY 460 TCAGAGAACTGCTCAATAT 440

DB 42 SerGluSncyspneAsnTyr 48

RESULT 33

O9VYNO PRELIMINARY; PRT: 189 AA.

AC O9VYNO: 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
GN CG15927 protein.

GN CG15927.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA April J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,

RA Baller R.M., Basu A., Bayendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari P., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,

RA Jajani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Mishina N.V., Mobery C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,

RA Palazolo M., Pittman G.S., Pan S., Polard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spler E., Spradell A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,

Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003489; AAF48162.1; -;
 DR FlyBase: FBgn0030389; CG15927.
 SQ SEQUENCE 189 AA: 22107 MW: 823E42FAF613FABF CRC64;

Alignment Scores:

Pred. No.:	428	Length:	189
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	5	Gaps:	0

US-09-817-318-1 (1-780) x Q9VYN0 (1-189)

OY 769 TTGCGTTTACACCAATATTTA 749
 |||||
 Db 160 LeuProLeuHisGlnTyrLeu 166

RESULT 34

Q8WCUS PRELIMINARY: PRT: 193 AA.
 AC Q8WCUS; 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE NADH dehydrogenase subunit II (Fragment).
 GN ND2.
 OS Rhodophis nuchalis.
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidodermata; Squamata; Scleroglossa; Serpentes; Colubroidae;
 OC Colubridae; Natricinae; Rhadophis.
 OX NCBI_TaxID=126483;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21617617; PubMed=11741383;
 RA Alfaro M.E., Arnold S.J.;
 RT "Molecular systematics and evolution of Regina and the thamnophiline snakes.";
 RL Mol. Phylogenet. Evol. 21:408-423(2001).
 DR EMBL: AF384826; AAL69655.1; -;
 DR InterPro: IPR001750; Oxidored_q1.
 DR Pfam: PF00361; oxidored_q1; 1.
 KW Mitochondrion.
 FT NON_TER 1
 FT NON_TER 193
 SQ SEQUENCE 193 AA: 21356 MW: 0CD9FD6D73AF929 CRC64;

Alignment Scores:

Pred. No.:	427	Length:	193
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	8	Gaps:	0

US-09-817-318-1 (1-780) x Q8WCUS (1-193)

OY 649 AACATACTTGTATTGACA 629
 |||||
 Db 145 AsnIleThrLeuLeuThr 151

RESULT 35

Q36064 PRELIMINARY: PRT: 196 AA.
 AC Q36064; 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Cytochrome b (Fragment).
 GN CYTB.

OS Trichurus lepturus.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopteleostei; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Scombroidei;
 OC Trichuridae; Trichurus.
 OX NCBI_TaxID=13733;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Block B.A., Finerty J.R., Stewart A.F.R., Kidd J.;
 RT "Evolution of endothermy in fish: Mapping physiological traits on a molecular phylogeny.";
 RL Science 0:0-0(1993).
 CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY BOUND TO THE PROTEIN (BY SIMILARITY).
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B, CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
 CC EMBL: L1154; AAA32130.1; -;
 DR InterPro: IPR00179; Cyt.b.b6.
 DR Pfam: PF00033; cytochrome_b_n; 1.
 DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
 KW Electron transport; Heme; Mitochondrion; Respiratory chain;
 FT NON_TER 1
 FT NON_TER 196
 SQ SEQUENCE 196 AA: 22093 MW: 999B30BC62F64930 CRC64;

Alignment Scores:

Pred. No.:	425	Length:	196
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	8	Gaps:	0

US-09-817-318-1 (1-780) x Q36064 (1-196)

OY 780 TTTTGTGTTTGGCTTTTAC 760
 |||||
 Db 44 PhePhePheCysLeuTyr 50

RESULT 36

Q8R9K7 PRELIMINARY: PRT: 197 AA.
 AC Q8R9K7; 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Predicted enzyme with a TIM-barrel fold.
 GN TTE1600.
 OS Thermomicrobacter tengcongensis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Thermomicrobacteriales; Thermomicrobacteriaceae; Thermomicrobacter.
 OX NCBI_TaxID=119072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MB4T / JCM11007;
 RX MEDLINE=21992816; PubMed=11997336;
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;
 RT "A complete sequence of *T. tengcongensis* genome.";
 RL Genome Res. 12:689-700(2002).
 DR EMBL: AE013115; AAM24804.1; -;
 KW complete proteome.
 SQ SEQUENCE 197 AA: 22642 MW: AF9FA99A01353CCD CRC64;

hypocretin-like
DKFZP434D146

```
Query Match: 2.868
```

Models: 0

```
US-09-817-318-1 (1-780) x Q9WK24 (1-219)
OY 760 GTAAAGCAAAAAAAAAAAAAA 780
    |||
Db 164 VALLYSALALYSLSLSLS 170

RESULT 40
Q9WK24 PRELIMINARY: PRT: 220 AA.
AC Q9WK24;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Reverse transcriptase (Fragment).
GN POL.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=260PT15W8C;
RX MEDLINE=20145414; PubMed=10681363;
RA Demeter L.M., Shafer R.W., Meehan P.M., Holden-Wiltse J., Fischl M.A.,
RA Freimuth W.W., Para M.F., Reichman R.C.;
RT "Delavirdine susceptibility and associated reverse transcriptase
RT mutations in human immunodeficiency virus type 1 isolates from
RT patients in a phase I/II trial of delavirdine monotherapy (ACTG
RT 260).";
RL Antimicrob. Agents Chemother. 44:794-797(2000).
DR EMBL: AF090478; AAC80472.1; -.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00078; rvt; 1.
KW RNA-directed DNA polymerase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 220 AA; 25790 MW; EDFDC39B0721EC12 CRC64;

Alignment Scores:
Pred. No.: 418 Length: 220
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
Gaps: 0

US-09-817-318-1 (1-780) x Q9WK24 (1-220)
OY 135 TATGAACATTTGGATATAA 115
    |||
Db 207 TYRGLUHLHLSTLSTLSTL 213

RESULT 41
Q9WK23 PRELIMINARY: PRT: 220 AA.
AC Q9WK23;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Reverse transcriptase (Fragment).
GN POL.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=260PT15W8P;
RX MEDLINE=20145414; PubMed=10681363;
RA Demeter L.M., Shafer R.W., Meehan P.M., Holden-Wiltse J., Fischl M.A.,
RA Freimuth W.W., Para M.F., Reichman R.C.;
RT "Delavirdine susceptibility and associated reverse transcriptase
RT mutations in human immunodeficiency virus type 1 isolates from
RT patients in a phase I/II trial of delavirdine monotherapy (ACTG
```

```
RT 260).";
RL Antimicrob. Agents Chemother. 44:794-797(2000).
DR EMBL: AF090479; AAC80473.1; -.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00078; rvt; 1.
KW RNA-directed DNA polymerase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 220 AA; 25753 MW; 770EB197D9F232E CRC64;

Alignment Scores:
Pred. No.: 418 Length: 220
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
Gaps: 0

US-09-817-318-1 (1-780) x Q9WK23 (1-220)
OY 135 TATGAACATTTGGATATAA 115
    |||
Db 207 TYRGLUHLHLSTLSTLSTL 213

RESULT 42
Q9WKX1 PRELIMINARY: PRT: 220 AA.
AC Q9WKX1;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Reverse transcriptase (Fragment).
GN POL.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=260PT26W8P;
RX MEDLINE=20145414; PubMed=10681363;
RA Demeter L.M., Shafer R.W., Meehan P.M., Holden-Wiltse J., Fischl M.A.,
RA Freimuth W.W., Para M.F., Reichman R.C.;
RT "Delavirdine susceptibility and associated reverse transcriptase
RT mutations in human immunodeficiency virus type 1 isolates from
RT patients in a phase I/II trial of delavirdine monotherapy (ACTG
RT 260).";
RL Antimicrob. Agents Chemother. 44:794-797(2000).
DR EMBL: AF090506; AAC80500.1; -.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00078; rvt; 1.
KW RNA-directed DNA polymerase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 220 AA; 25750 MW; BC161482A7AZCE0A CRC64;

Alignment Scores:
Pred. No.: 418 Length: 220
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
Gaps: 0

US-09-817-318-1 (1-780) x Q9WKX1 (1-220)
OY 135 TATGAACATTTGGATATAA 115
    |||
Db 207 TYRGLUHLHLSTLSTLSTL 213

RESULT 43
Q9WKW7 PRELIMINARY: PRT: 220 AA.
AC Q9WKW7;
```

DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Reverse transcriptase (Fragment).
OS POL.
GN Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBI_TaxID=11676;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=260P28W8C;
RA MEDLINE=2015414; PubMed=10681363;
RA Demeter L.M., Shaler R.W., Meenan P.M., Holden-Wiltse J., Fischl M.A.,
RA Freimuth W.W., Para M.F., Reichman R.C.;
RT "Delavirdine susceptibilities and associated reverse transcriptase
RT mutations in human immunodeficiency virus type 1 isolates from
RT patients in a phase I/II trial of delavirdine monotherapy (ACTG
RT 260).";
RL Antimicrob. Agents Chemother. 44:794-797(2000).
DR EMBL; AF090510; AAC80504.1; -;
DR InterPro: IPR000477; RVTse.
DR Pfam: PF000078; rvt.1.
KW RNA-directed DNA polymerase.
FT NON_TER 1
FT NON_TER 220
SQ SEQUENCE 220 AA; 25751 MW; 56CDA18384B026B8 CRC64;

Alignment Scores:

Pred. No.:	418	Length:	220
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	15	Gaps:	0

US-09-817-318-1 (1-780) x Q9KFW7 (1-220)

QY 135 TATGAAGTACATTTGGATAAA 115

Db 207 TyrgluuhtslsleuAsplys 213

RESULT 44

Q8XFN8 PRELIMINARY; PRT; 222 AA.
ID Q8XFN8;
AC Q8XFN8;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE GTP cyclohydrolase I (EC 3.5.4.16).
GN FOIE OR STM2193 OR STY2427.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
NCBI_TaxID=602, 601;
RN (1)
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SSGC1412 / ATCC 700720;
RA MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porciliak S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RA MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,

RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Reltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen J.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [1]
RP SEQUENCE FROM N.A.
RC EMBL; AF008797; AA121097.1; -;
DR EMBL; AL627273; CAD02575.1; -;
DR InterPro: IPR001474; GTP_cyclohydrol.
DR Pfam: PF01227; GTP_cyclohydrol.1.
DR ProDom: PD003330; GTP_cyclohydrol.1.
DR TIGRFAMS: TIGR00063; foie.1.
DR PROSITE: PS00859; GTP_CYCLOHYDROL_1_1; 1.
DR PROSITE: PS00860; GTP_CYCLOHYDROL_1_2; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 222 AA; 24770 MW; 906EB3877D1F444E CRC64;

Alignment Scores:

Pred. No.:	417	Length:	222
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.86%	Indels:	0
DB:	16	Gaps:	0

US-09-817-318-1 (1-780) x Q8XFN8 (1-222)

QY 233 ACAGCTTACAGACCTTACTC 253

Db 161 Thralaenclnhrleu 167

RESULT 45

Q99VW2 PRELIMINARY; PRT; 224 AA.
ID Q99VW2;
AC Q99VW2;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical protein SA0614.
GN SA0614.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus;
NCBI_TaxID=158679;
RN (1)
RP SEQUENCE FROM N.A.
RC MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Iian J., Ito T., Kanamori M.,
RA Matsunaru H., Maruyama A., Murakami H., Hoshiyama A., Mizutani Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hirakawa H., Kuhara S., Goto S., Yabuuchi J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
CC -!- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSOR TRANSDUCTION SYSTEMS.
CC -!- SIMILARITY: TO BACTERIAL REGULATORY PROTEINS INVOLVED IN SIGNAL
CC TRANSDUCTION.
DR EMBL; AP003131; BAB41847.1; -;
DR HSSP: P08402; 1B00.
DR InterPro: IPR001789; Response_reg.
DR InterPro: IPR001867; Trans_reg_C.
DR Pfam: PF00072; response_reg.1.
DR Pfam: PF00486; trans_reg_C.1.
DR ProDom: PD000039; Response_reg.1.
DR ProDom: PD000329; Trans_reg_C.1.
DR SMART: SM00448; REC; 1.

KW DNA-binding; Phosphorylation; Sensory transduction;
KW Transcription regulation; Hypothetical protein; Complete proteome.
SQ SEQUENCE 224 AA; 26066 MW; 4A636CBAA6FAD233 CRC64;

Alignment Scores:

Pred. No.:	416	Length:	224
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	16	Gaps:	0

US-09-817-318-1 (1-780) x Q99VW2 (1-224)

QY 542 ACAGTCAATGTAATAGATTG 522

DB 193 ThrValAsnValAsnArgLeu 199

Search completed: February 4, 2003, 07:20:21
Job time : 67.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 4, 2003, 07:17:29; Search time 43.5 seconds

(Without alignments)
4778.643 Million cell updates/sec

Title: US-09-817-318-1

Perfect score: 245
Sequence: 1 aactactgtctctcttaag.....taaggcaaaaaaaaaa 780

Scoring table:

OLIGO
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 908470 segs, 133250620 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1687582

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=framer_n2p_model -DEV=xlp
-Q=cg92_l1/USPTO.spool/US09817318/runat_04022003_071721_6604/app_query.fasta.1.967
-DB=A.Geneseq.101002 -OPM=fastan -SOFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITs=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=45 -MODE=LOCAL
-OUTFM=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USR=US09817318.@CGN_L1_30_@runat_04022003_071721_6604 -NCPU=6 -ICPU=3
-NO_XLIFY -NO_MMAP -LARGEOUT -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-MAIN_TIMEOUT=30 -THREADS=1 -KAPOP=60 -XGAPEXT=60 -Fgapop=6 -Fgapext=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

A.Geneseq.101002: *
1: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA2001.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA2002.DAT: *

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	8	3.3	51	22	ABG20710	Novel human diagno
2	8	3.3	55	22	ABG07001	Novel human diagno
3	8	3.3	93	22	AAAB7296	Human immune/haema
4	8	3.3	341	23	ABAB2615	Heridically activ
5	8	3.3	397	21	AAAG29847	Arabidopsis thalia
6	8	3.3	397	21	AAAG34905	Arabidopsis thalia
7	8	3.3	454	21	AAAG29846	Arabidopsis thalia
8	8	3.3	454	21	AAAG34904	Arabidopsis thalia
9	8	3.3	455	21	AAAG29845	Arabidopsis thalia
10	8	3.3	470	21	AAAG34903	Arabidopsis thalia
11	8	3.3	593	21	AAAG23149	Arabidopsis thalia
12	8	3.3	593	21	AAAG60457	Arabidopsis thalia
13	8	3.3	645	21	AAAG23148	Arabidopsis thalia
14	8	3.3	645	21	AAAG60456	Arabidopsis thalia
15	8	3.3	716	21	AAAG60455	Arabidopsis thalia
16	8	3.3	717	21	AAAG23147	Arabidopsis thalia
17	7	2.9	21	16	AAAG4158	Peptide enhancer o
18	7	2.9	21	16	AAAG77072	Synthetic anti-neo
19	7	2.9	21	16	AAAG74721	Tryptic digestion
20	7	2.9	21	16	AAAG4800	Amphipathic peptid
21	7	2.9	21	17	AAAG2402	Lytic peptide used
22	7	2.9	21	17	AAAG90756	Synthetic lytic pe
23	7	2.9	21	19	AAAG39773	Synthetic lysine r
24	7	2.9	28	22	ABG27285	Novel human diagno
25	7	2.9	38	21	AAAG68253	Murine class I mol
26	7	2.9	38	21	AAAG52907	Murine class I mol
27	7	2.9	38	22	AAAB58668	Murine class I pr
28	7	2.9	42	21	AAAG68254	Murine class I mol
29	7	2.9	42	21	AAAG52908	Murine class I mol
30	7	2.9	42	22	AAAG7287	Human immune/haema
31	7	2.9	42	22	AAAB58669	Murine class I pr
32	7	2.9	46	22	ABG01742	Novel human diagno
33	7	2.9	71	22	AAAB63839	Human prostate can
34	7	2.9	71	22	AAAB63843	Human prostate can
35	7	2.9	87	17	AAAG07875	Protein reactive w
36	7	2.9	87	21	AAAG03475	Human secreted pro
37	7	2.9	88	20	AAAG12861	Human 5' EST seque
38	7	2.9	88	22	AAAG07086	Human polypeptide
39	7	2.9	91	23	ABP03162	Human ORFX protein
40	7	2.9	94	21	AAAG25318	Zea mays protein f
41	7	2.9	97	23	ABAB80897	Murine ectoxin pol
42	7	2.9	98	23	ABAB80898	Murine ectoxin pol
43	7	2.9	99	22	AAAG59212	Protonibacterium
44	7	2.9	100	22	AAAG07453	Human polypeptide
45	7	2.9	104	23	ABP31587	Human ORF560 prote

ALIGNMENTS

RESULT 1
ID ABG20710 standard; Protein: 51 AA.

AC ABG20710;
DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #20701.

KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

XX 11-OCT-2001.
XX PD

PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
DR N-PSDB: AAS84897.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
PS Claim 20; SEQ ID No 51069; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 51 AA;
XX
XX Alignment Scores:
Pred. No.: 29.2 Length: 51
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 22 Gaps: 0
XX
US-09-817-318-1 (1-780) x ABG20710 (1-51)
OY 448 AGCAGTCTCTGAAATCACCAGAA 471
DB 33 SerSerSerLeuYSerProLys 40
XX
RESULT 2
ABG07001
ID ABG07001 standard; Protein: 55 AA.
XX
AC ABG07001;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #6992.
XX
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX

XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
PF
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
DR N-PSDB: AAS71188.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
PS Claim 20; SEQ ID No 37360; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 55 AA;
XX
XX Alignment Scores:
Pred. No.: 29.1 Length: 55
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 22 Gaps: 0
XX
US-09-817-318-1 (1-780) x ABG07001 (1-55)
OY 193 AAATTAACATTTATTCATTTT 170
DB 29 LysLeuThrPheYSerPhePhe 36
XX
RESULT 3
AAM87296
ID AAM87296 standard; Protein: 93 AA.
XX
AC AAM87296;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen SEQ ID NO:14889.
XX
XX Human: immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.
XX
XX

XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI: 2001-483426/52.
DR N-PSDB: AAK60077.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Claim 11: SEQ ID NO 14889; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK67694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 93 AA;

Alignment Scores:
Pred. No.: 28 Length: 93
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 22 Gaps: 0

US-09-817-318-1 (1-780) x AAM87296 (1-93)
OY 634 TTGACAGTACTTGTGATTATCC 611
DB 71 LeuThValThrSerAspLeuSer 78

RESULT 4
AB92615
ID ABB92615 standard; Protein; 341 AA.
XX
AC ABB92615;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 1826.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN MO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB) BAYER AG.
XX
PI Tiejfen K, Weidler M;
XX
DR WPI: 2002-269010/31.

XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 1826; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
SQ Sequence 341 AA;

Alignment Scores:
Pred. No.: 25.7 Length: 341
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 23 Gaps: 0

US-09-817-318-1 (1-780) x ABB92615 (1-341)
OY 98 ACTAGTTTATTTTATATATATCC 121
DB 3 SerLysPheIleLeuTyrLeuSer 10

RESULT 5
AAG29847
ID AAG29847 standard; Protein; 397 AA.
XX
AC AAG29847;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35580.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP103405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.:	25.4	Length:	397
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.27%	Indels:	0
DB:	21	Gaps:	0

US-09-817-318-1 (1-780) x AAG29847 (1-397)

QY 587 AAAAGATATTAGATATATGGCAG 610

Db 266 LysIleuAspAsnGlyCln 273

RESULT 6

ID AAG34905 standard; Protein: 397 AA.

XX AAG34905;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 42549.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0136299.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142399.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144864.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.

PR	22-JUL-1999;	99US-014089.
PR	22-JUL-1999;	99US-0145192.
PR	22-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147312.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0155139.
PR	22-SEP-1999;	99US-0155486.
PR	23-SEP-1999;	99US-0155659.
PR	24-SEP-1999;	99US-0156458.
PR	28-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159350.
PR	14-OCT-1999;	99US-0159351.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Alignment Scores:		
Pred. No.:	25.4	Length: 397
Score:	8.00	Matches: 8
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	3.27%	Indels: 0
DB:	21	Gaps: 0
US-09-817-318-1 (1-780) x AAG34905 (1-397)		
Oy	587	AAAGATATTAGTAAATGGCAG 610
Db	266	Lysdysileuakspasnglygn 273
RESULT 7		
AAG29846		
ID	AAG29846 standard; Protein: 454 AA.	
XX	AAG29846;	
XX	17-OCT-2000 (first entry)	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 35579.	
XX	Protein identification:	

PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134270.
PR 18-MAY-1999; 9905-0134768.
PR 18-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138540.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144684.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.

PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148365.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151080.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.

```
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 25.2
Score: 8.00
Percent Similarity: 100.00%
Best local Similarity: 100.00%
Query Match: 3.27%
DB: 21
Gaps: 0

US-09-817-318-1 (1-780) x AAG29846 (1-454)

OY 587 AAAAGATTTAGTAAATGGCAG 610
    |||||
Db 323 LysLysIleuaspasnGlyGln 330

RESULT 8
AAC34904
ID AAG34904 standard; Protein; 454 AA.
XX
AC AAG34904;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 42548.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
```

```
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135253.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
```

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149802.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 25.2
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.27%
DB: 21
Gaps: 0

US-09-817-318-1 (1-780) x AAG34904 (1-454)

OY 587 AAAAGATATTAGATTAATGGCAG 610
DB 323 LysLysIleuaspasnGlycin 330

RESULT 9
AAG29845
ID AAG29845 standard; Protein; 455 AA.
XX
AC AAG29845;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35578.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159285.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

```
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 25 2
Score: 8.00 Length: 455
Percent Similarity: 100.00% Matches: 8
Best Local Similarity: 100.00% Conservative: 0
Query Match: 3.27% Mismatches: 0
DB: 21 Indels: 0
Gaps: 0

US-09-817-318-1 (1-780) x AAG29845 (1-455)
QY 587 AAAAGATATTAGTAAATGCGCAG 610
Db 324 LysylsilleuuspaspnnglyGln 331

RESULT 10
AAG34903
ID AAG34903 standard; Protein; 470 AA.
AC AAG34903;
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 42547.
DE
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
PN EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0127462.
PR 01-APR-1999; 99US-0128234.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 30-APR-1999; 99US-0132484.
PR 04-MAY-1999; 99US-0132485.
PR 05-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
```

```
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139464.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
```

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. NO.: 25.1
Score: 8.00
Percent Similarity: 100.00%
Best local Similarity: 100.00%
Query Match: 3.27%
DB: 21
Gaps: 0

US-09-817-318-1 (1-780) x AAG34903 (1-470)

QY 587 AAAAGATATAGATTAATGGCAG 610
|||||
Db 339 LysLysIleLeuAspAsnGlyGln 346

RESULT 11
AAG23149
ID AAG23149 standard; Protein: 593 AA.
XX
AC AAG23149;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 26348.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
FN EP1033405-A2.
XX
PD 06-SEP-2000.
XX

PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135529.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147792.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158023.
PR 08-OCT-1999; 99US-0158233.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161922.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

Alignment Scores:

Pred. No.:	24.7	Length:	593
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.33%	Indels:	0
DB:	21	Gaps:	0

US-09-817-318-1 (1-780) x AAG23149 (1-593)

QY 482 CATTAATCAGTTGGTGATTTC 459

Db 398 HislySileSerLeuGlyAspPhe 405

RESULT 12

AAG60457
ID AAG60457 standard; Protein: 593 AA.

AC AAG60457;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 78307.

KW Protein identification: signal transduction pathway; metabolic pathway;

KM hybridisation assay; genetic mapping; gene expression control; promoter;

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 22-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 28-APR-1999; 990S-0130891.
PR 30-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 06-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134321.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.

PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 17-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 24-JUN-1999; 990S-0140354.
PR 28-JUN-1999; 990S-0140695.
PR 29-JUN-1999; 990S-0140823.
PR 30-JUN-1999; 990S-0140991.
PR 01-JUL-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142380.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 24.7
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.33%
DB: 21
Gaps: 0

US-09-817-318-1 (1-780) x AAG60457 (1-593)

QY 482 CATAAATCAGTTGGTGATTTTC 459
|||||
Db 398 HisLysIleSerLeuGlyAspHe 405

RESULT 13
AAG23148
ID AAG23148 standard; Protein; 645 AA.
XX
AC AAG23148;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 26347.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132487.
PR 04-MAY-1999; 99US-0132487.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155466.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158222.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159299.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

PR 05-AUG-1999: 99US-0147192.
PR 06-AUG-1999: 99US-0147260.
PR 06-AUG-1999: 99US-0147303.
PR 06-AUG-1999: 99US-0147416.
PR 09-AUG-1999: 99US-0147493.
PR 09-AUG-1999: 99US-0147935.
PR 10-AUG-1999: 99US-0148171.
PR 11-AUG-1999: 99US-0148319.
PR 12-AUG-1999: 99US-0148341.
PR 13-AUG-1999: 99US-0148565.
PR 16-AUG-1999: 99US-0148684.
PR 17-AUG-1999: 99US-0149368.
PR 18-AUG-1999: 99US-0149175.
PR 18-AUG-1999: 99US-0149426.
PR 20-AUG-1999: 99US-0149722.
PR 20-AUG-1999: 99US-0149723.
PR 20-AUG-1999: 99US-0149929.
PR 23-AUG-1999: 99US-0149902.
PR 23-AUG-1999: 99US-0149930.
PR 25-AUG-1999: 99US-0150566.
PR 26-AUG-1999: 99US-0150884.
PR 27-AUG-1999: 99US-0151065.
PR 27-AUG-1999: 99US-0151066.
PR 27-AUG-1999: 99US-0151080.
PR 30-AUG-1999: 99US-0151303.
PR 31-AUG-1999: 99US-0151438.
PR 01-SEP-1999: 99US-0151930.
PR 07-SEP-1999: 99US-0152363.
PR 10-SEP-1999: 99US-0153070.
PR 13-SEP-1999: 99US-0153758.
PR 15-SEP-1999: 99US-0154018.
PR 16-SEP-1999: 99US-0154039.
PR 20-SEP-1999: 99US-0154779.
PR 22-SEP-1999: 99US-0155139.
PR 23-SEP-1999: 99US-0155486.
PR 24-SEP-1999: 99US-0155659.
PR 28-SEP-1999: 99US-0156458.
PR 29-SEP-1999: 99US-0156596.
PR 04-OCT-1999: 99US-0157117.
PR 05-OCT-1999: 99US-0157753.
PR 06-OCT-1999: 99US-0157865.
PR 07-OCT-1999: 99US-0158029.
PR 08-OCT-1999: 99US-0158232.
PR 12-OCT-1999: 99US-0158369.
PR 13-OCT-1999: 99US-0159293.
PR 13-OCT-1999: 99US-0159294.
PR 13-OCT-1999: 99US-0159295.
PR 14-OCT-1999: 99US-0159329.
PR 14-OCT-1999: 99US-0159330.
PR 14-OCT-1999: 99US-0159331.
PR 14-OCT-1999: 99US-0159637.
PR 14-OCT-1999: 99US-0159638.
PR 18-OCT-1999: 99US-0159584.
PR 21-OCT-1999: 99US-0160741.
PR 21-OCT-1999: 99US-0160767.
PR 21-OCT-1999: 99US-0160768.
PR 21-OCT-1999: 99US-0160770.
PR 21-OCT-1999: 99US-0160814.
PR 21-OCT-1999: 99US-0160815.
PR 22-OCT-1999: 99US-0160981.
PR 22-OCT-1999: 99US-0160989.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 26-OCT-1999: 99US-0161359.
PR 26-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161920.
PR 28-OCT-1999: 99US-0161992.
PR 28-OCT-1999: 99US-0161993.
PR 29-OCT-1999: 99US-0162142.

Alignment Scores:
Pred. No.: 24 6
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.33%
DB: 21
Gaps: 0
US-09-817-318-1 (1-780) x AAG60456 (1-645)
QY 482 CATAAATCAGTTGGTGATTTC 459
|||||
Db 450 HisLysIleSerLeuGlyAspPhe 457
RESULT 15
AAG60455
ID AAG60455 standard; Protein; 716 AA.
XX
AC AAG60455.
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 78305.
XX
KW Protein identification: signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999: 99US-0121825.
PR 05-MAR-1999: 99US-0123180.
PR 09-MAR-1999: 99US-0123548.
PR 23-MAR-1999: 99US-0125788.
PR 25-MAR-1999: 99US-0126264.
PR 29-MAR-1999: 99US-0126785.
PR 01-APR-1999: 99US-0127462.
PR 06-APR-1999: 99US-0128234.
PR 08-APR-1999: 99US-0128714.
PR 16-APR-1999: 99US-0129845.
PR 19-APR-1999: 99US-0130077.
PR 21-APR-1999: 99US-0130449.
PR 23-APR-1999: 99US-0130510.
PR 23-APR-1999: 99US-0130891.
PR 28-APR-1999: 99US-0131449.
PR 30-APR-1999: 99US-0132048.
PR 30-APR-1999: 99US-0132407.
PR 04-MAY-1999: 99US-0132484.
PR 05-MAY-1999: 99US-0132485.
PR 06-MAY-1999: 99US-0132486.
PR 06-MAY-1999: 99US-0132487.
PR 07-MAY-1999: 99US-0132863.
PR 11-MAY-1999: 99US-0134256.
PR 14-MAY-1999: 99US-0134218.
PR 14-MAY-1999: 99US-0134219.
PR 14-MAY-1999: 99US-0134221.
PR 14-MAY-1999: 99US-0134570.
PR 18-MAY-1999: 99US-0134768.
PR 18-MAY-1999: 99US-0134941.
PR 19-MAY-1999: 99US-0135124.
PR 20-MAY-1999: 99US-0135124.
PR 21-MAY-1999: 99US-0135533.
PR 24-MAY-1999: 99US-0135629.
PR 25-MAY-1999: 99US-0136021.
PR 27-MAY-1999: 99US-0136392.
PR 28-MAY-1999: 99US-0136782.
PR 01-JUN-1999: 99US-0137222.
PR 03-JUN-1999: 99US-0137528.

PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0138119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 17-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145911.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.

PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149175.
PR 20-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151066.
PR 30-AUG-1999; 990S-0151080.
PR 31-AUG-1999; 990S-0151303.
PR 01-SEP-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161992.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

Alignment Scores:
Pred. No.:

24.4

Length:

716

Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 21 Gaps: 0

US-09-817-318-1 (1-780) x AAG60455 (1-716)

OY 482 CATAAATCAGTTGGTGATTTC 459
|||||
Db 521 HistlystleSerLeuGlyAspPhe 528

RESULT 16
AAG23147
ID AAG23147 standard; Protein: 717 AA.
XX AAG23147;
AC AAG23147;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 26346.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 9905-0121825.
PR 05-MAR-1999; 9905-0123180.
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126284.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0129845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130449.
PR 23-APR-1999; 9905-0130510.
PR 28-APR-1999; 9905-0130891.
PR 30-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 04-MAY-1999; 9905-0132407.
PR 05-MAY-1999; 9905-0132484.
PR 06-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 07-MAY-1999; 9905-0132487.
PR 11-MAY-1999; 9905-0132483.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134376.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136021.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0136782.
PR 03-JUN-1999; 9905-0137222.
PR 04-JUN-1999; 9905-0137528.
PR 07-JUN-1999; 9905-0137502.
PR 9905-0137724.

PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139452.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 21-JUN-1999; 9905-0139763.
PR 22-JUN-1999; 9905-0139817.
PR 23-JUN-1999; 9905-0139859.
PR 23-JUN-1999; 9905-0140353.
PR 24-JUN-1999; 9905-0140354.
PR 28-JUN-1999; 9905-0140695.
PR 29-JUN-1999; 9905-0140823.
PR 30-JUN-1999; 9905-0140991.
PR 01-JUL-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 02-JUL-1999; 9905-0142154.
PR 06-JUL-1999; 9905-0142055.
PR 08-JUL-1999; 9905-0142390.
PR 09-JUL-1999; 9905-0142803.
PR 12-JUL-1999; 9905-0142920.
PR 13-JUL-1999; 9905-0142977.
PR 14-JUL-1999; 9905-0143542.
PR 15-JUL-1999; 9905-0143624.
PR 16-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 26-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0146389.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.

```

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157863.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:

24.4
8.00
100.00%

Length:
Matches:
Conservative:

717
8
0

```

Best Local Similarity: 100.00%
Query Match: 3.33%
DB: 21
Mismatches: 0
Indels: 0
Gaps: 0

US-09-817-318-1 (1-780) x AAR84158 (1-717)

QY 482 CATMAATCAGTTGGGTGATTTTC 459
|||||
Db 522 HSLYSILSERLEUGLYASPHE 529

RESULT 17
AAR84158
ID AAR84158 standard; peptide; 21 AA.
XX
AC AAR84158;
XX
DT 06-JUN-1996 (first entry)
XX
DE Peptide enhancer of fibroblast and keratinocyte proliferation.
XX
KW amphipathic peptide; enhance; fibroblast; keratinocyte; proliferation;
KW wound healing; defensin; antimicrobial.
XX
OS Synthetic.
XX
PN WO9528832-A1.
XX
PD 02-NOV-1995.
XX
PF 19-APR-1995; 95WO-US04718.
XX
PR 20-APR-1994; 94US-0231730.
XX
PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.
XX
PI Jaynes JM, Julian GR;
XX
DR WPI: 1995-382791/49.
XX
PT Use of amphipathic peptide(s) - for enhancing fibroblast and
PT keratinocyte proliferation to promote wound healing in a mammal
XX
XX
PS Claim 7; Page 52-53; 64pp; English.
XX
CC AAR84128-73 are amphipathic peptides which are able to stimulate the
CC proliferative growth of fibroblasts and epithelial cells such as
CC keratinocytes, hence enhancing wound healing in mammalian subjects.
CC The peptides concomitantly have antimicrobial efficacy, against
CC microbial species including those which cause or otherwise mediate
CC sepsis and wound infection.
XX
SQ Sequence 21 AA;

Alignment Scores:
Pred. No.: 325
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.86%
DB: 16
Mismatches: 0
Indels: 0
Gaps: 0

US-09-817-318-1 (1-780) x AAR84158 (1-21)

QY 760 GTPAAGCAAAAAAAAAAAAAA 780
|||||
Db 15 VALYSALALYSLSYLSYS 21

RESULT 18
AAR7072
ID AAR7072 standard; peptide; 21 AA.
XX
AC AAR7072;
XX

```

```

DT 20-MAY-1996 (first entry)
XX
XX Synthetic anti-neoplastic lytic peptide.
DE
XX Anti-cancer; lysis; amphipathic; neoplasia; tumour; cystic fibrosis;
XX bronchopulmonary; viral; virus; analogue; magainin; cecropin;
KW melittin; defensin.
XX
XX Synthetic.
OS
XX
XX W09527497-A1.
XX
XX 19-OCT-1995.
XX
XX 06-APR-1995; 95WO-US04335.
XX
XX 08-APR-1994; 94US-0225476.
XX
XX (DEME-) DEMETER BIOTECHNOLOGIES LTD.
XX
XX Jaynes JM, Julian GR:
XX
XX WPI: 1995-366226/47.
XX
XX Method of combatting mammalian neoplasias and other disease states -
PT by delivering non-naturally occurring, non-onco-cytologically
PT proliferative lytic peptide
PT
PS Claim 7; Page 50; 63pp; English.
XX
XX AAR7042-R77081 are synthetic, amphipathic, lytic peptide analogues of
XX melittin, cecropin, magainin and defensin peptides. The peptides are
XX between 23 and 39 residues long, are amphipathic, carry an overall
XX positive charge and have anti-neoplastic activity. The peptides are
XX specifically useful for the lysis of cancer cells. Normal mammalian
XX cells are resistant to lysis due to their highly organised
XX cytoskeleton, cancerous cells however possess an inferior and
XX structurally compromised cytoskeleton which when acted upon by lytic
XX peptides will cause cell lysis. This allows the lytic peptides to be
XX used for in vivo treatment of cancers. The peptides are esp. useful
XX for the treatment of female mammalian cancers e.g. breast, ovarian,
XX uterine and cervical cancers. The peptides can however be used to
XX treat most forms of cancer, cystic fibrosis, pneumonia, bronchitis,
XX and bronchopulmonary viral and microbial infections.
XX
XX Sequence 21 AA;
SQ
Alignment Scores:
Pred. No.: 325
Score: 7.00 Length: 21
Percent Similarity: 100.00% Matches: 7
Best Local Similarity: 100.00% Conservative: 0
Query Match: 2.86% Mismatches: 0
DB: 16 Indels: 0
Gaps: 0
US-09-817-318-1 (1-780) x AAR77072 (1-21)
QY 760 GTAAAGCGCAAAAAAAAAAAAAA 780
Db 15 VallysAlaLysLysLysLys 21
RESULT 19
AAR74721
ID AAR74721 standard; peptide; 21 AA.
XX
XX AAR74721;
XX
XX 11-JAN-1996 (first entry)
DE Tryptic digestion resistant lytic peptide.
XX
XX Lytic; methylated; tryptic resistant; immunological; infection;
KW neoplasia.
XX

```

OS	XX	Synthetic.	Location/Qualifiers
XX	XX	Key	1..21
XX	XX	Modified-site	/note=
XX	XX	FT	"the epsilon-amino groups of the lysine
XX	XX	FT	residues and the alpha-amino group of the
XX	XX	FT	N-terminal amino acid are sufficiently
XX	XX	FT	methylated to impart enhanced proteolytic
XX	XX	FT	digestion resistance to the peptide"
XX	XX	PN	
XX	XX	W09513085-A1.	
XX	XX	18-MAY-1995.	
XX	XX	01-NOV-1994;	94WO-US12550.
XX	XX	08-NOV-1993;	93US-0148889.
XX	XX	(DEME-)	DEMEETER BIOTECHNOLOGIES LTD.
XX	XX	Julian GR;	
XX	XX	WPI: 1995-193900/25.	
XX	XX	New lytic peptide(s) for treating e.g. infections and neoplasias -	
XX	XX	PT contain mainly alanine, valine and lysine residues with the lysine	
XX	XX	residues pref. methylated	
XX	XX	Claim 8; Page 50; 66pp; English.	
XX	XX	AAR74699-R74733 are synthetic lytic peptides which are rich in the	
XX	XX	amino acids lysine, valine and alanine. They are useful in the	
XX	XX	treatment of neoplasia and viral, bacterial, protozoan, fungal or	
XX	XX	yeast infections. The lysine residues and the N-terminal amino acid	
XX	XX	residue of the peptides are methylated which imparts enhanced	
XX	XX	resistance to proteolytic digestion and allows effective in vivo	
XX	XX	administration.	
XX	XX	Sequence	21 AA;
XX	XX	Alignment Scores:	
XX	XX	Pred. No.:	325
XX	XX	Score:	7..00
XX	XX	Percent Similarity:	100.00%
XX	XX	Best local Similarity:	100.00%
XX	XX	Query Match:	2.86%
XX	XX	DB:	16
XX	XX	US-09-817-318-1 (1-780) x AAR74721 (1-21)	
XX	XX	OY	760 GTAAAGCAAAAAAAAAAAAAA 780
XX	XX	Db	
XX	XX	15 Valylsalalysylslyslys 21	
XX	XX	RESULT 20	
XX	XX	AAR64800	
XX	XX	ID AAR64800 standard; peptide; 21 AA.	
XX	XX	AC AAR64800;	
XX	XX	DT 24-AUG-1995 (first entry)	
XX	XX	DE Amphipathic peptide #31, for treating a pulmonary disease state.	
XX	XX	XX Amphipathic peptide; pulmonary disease; resistant; bacteria;	
XX	XX	KM proteolytic digestion; methylation; glyoxylation; cystic fibrosis;	
XX	XX	KW CF; neoplasia; pneumonia; bronchitis; lytic activity; lysis.	
XX	XX	OS Synthetic.	
XX	XX	Key	Location/Qualifiers
XX	XX	Modified-site	1

FT Modified-site /note- "Opt. methylated or glyoxylated"
FT 1...21 /note- "Lys residues are opt. methylated"
XX
XX
PN WO9428921-A.
XX
PD 22-DEC-1994.
XX
PF 02-JUN-1994; 94MO-US06176.
XX
PR 04-JUN-1993; 93US-0039620.
XX
PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.
XX
PI Jaynes JM, Julian GR;
XX
DR WPI; 1995-036106/05.
XX
PT Treatment of pulmonary disease states such as cystic fibrosis -
XX by admin. of a non-naturally occurring amphiphatic peptide
XX
PS Claim 21; Page 50; 54pp; English.
XX
CC The sequences given in AAR64770-806 and AAR76077 represent non-naturally
CC occurring amphiphatic peptides which may be used for treating
CC pulmonary disease states. These peptides have enhanced resistance
CC to proteolytic digestion due to methylation of the epsilon-amino
CC group of lys residues or glyoxylation of the side chains of Arg
CC residues, and methylation or glyoxylation of the N-terminal amino
CC acid. The peptides are prepared by standard methods of solid phase
CC synthesis and may be used in the treatment of cystic fibrosis (CF),
CC neoplasia, pneumonia, bronchitis, etc.. The peptides pref. have a
CC lytic activity, thereby lysing pathogenic bacteria, virally infected
CC cells and transformed cells as well as treating the epithelial cell
CC defect of CF.
XX
SQ Sequence 21 AA:

Alignment Scores:
Pred. No.: 325 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x AAR64800 (1-21)

QY 760 GTAAGCGCAAAAAAAAAA 780
DB 15 VallysAlaLysLysLysLys 21

RESULT 21
AAR92402
ID AAR92402 standard; peptide; 21 AA.
XX
AC AAR92402;
XX
DT 17-SEP-1996 (first entry)
XX
DE Lytic peptide used in ubiquitin-lytic peptide fusion protein.
XX
KW Ubiquitin; fusion protein; lysis; infection; neoplasia; wound healing;
XX stability; reduced toxicity.
XX
OS Synthetic.
XX
PN WO9603519-A1.
XX
PD 08-FEB-1996.
XX
PF 24-JUL-1995; 95WO-US09339.
XX

PR 22-JUL-1994; 94US-0279472.
XX
XX (DEME-) DEMETER BIOTECHNOLOGIES LTD.
PA (USDA) US SEC OF AGRIC.
XX
XX Belknap W, Garbarino J, Jaynes J;
PI WPI; 1996-117061/12.
XX
DR
XX
XX
PT New fusion protein of ubiquitin and a lytic peptide - for treating
PT infections and neoplasia, healing wounds, etc. also related nucleic
PT acid, vectors, and transformed cells
XX
XX
PS Claim 5; Page 21; 112pp; English.
XX
CC AAR92372-R92462 are lytic peptides used to create ubiquitin-lytic
CC peptide fusion proteins in which the ubiquitin polypeptide is linked
CC at its 3'-terminus to the lytic peptide. The lytic peptides are pref.
CC selected from either the cecropins, defensins, sarcotoxins, melittin
CC and magainins. The fusion proteins (FPS) are useful for treating
CC protozoal, bacterial, fungal and viral infections and neoplasia (in
CC plants and animals) in the same way as the FP alone, they also
CC promote wound healing. FPS produced in bacteria may be cleaved in
CC vitro by ubiquitin hydrolases to recover the active lytic peptide.
CC FPS produced in eukaryotic cells are cleaved by endogenous enzymes
CC to yield lytic peptide. Recombinant DNA encoding the FPS have
CC greater stability in bacteria than DNA encoding the lytic peptide
CC only.
XX
SQ Sequence 21 AA:

Alignment Scores:
Pred. No.: 325 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x AAR92402 (1-21)

QY 760 GTAAGCGCAAAAAAAAAA 780
DB 15 VallysAlaLysLysLysLys 21

RESULT 22
AAR90756
ID AAR90756 standard; peptide; 21 AA.
XX
AC AAR90756;
XX
DT 17-SEP-1996 (first entry)
XX
DE Synthetic lytic peptide #86.
XX
XX Lytic peptide; ubiquitin; synthetic analogue; cell membrane; cell lysis;
XX microbial pathogen; disease-resistant plant; bacterial infection; fungus;
XX protozoa; virus; neoplasia; fusion protein; hydrolase.
XX
OS Synthetic.
XX
PN WO9603522-A1.
XX
PD 08-FEB-1996.
XX
PF 24-JUL-1995; 95WO-US09338.
XX
PR 22-JUL-1994; 94US-0279472.
XX
PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.
XX
PI Jaynes J;
XX

DR WPI: 1996-117064/12.
XX Lytic peptide(s), useful for developing disease-resistant plants -
PT can be expressed as fusion protein with ubiquitin for stable prodn.
PT In bacterial host cells
XX
PS Example 1: Page 21; 11pp; English.
XX
CC AAR89967-R90021 and AAR90726-R90763 represent synthetic analogues of
CC naturally occurring lytic peptides. Lytic peptides destroy prokaryotic
CC and other non-host cells by disrupting the cell membrane and promoting
CC cell lysis. Synthetic lytic peptide analogues have similar or higher
CC levels of lytic activity for many different types of cells, compared to
CC naturally occurring forms. The concentration of the synthetic analogue
CC required to lyse microbial pathogens does not lyse normal mammalian
CC cells. The lytic peptides can be expressed in plants to allow for the
CC development of disease-resistant plants. The peptides are useful in
CC promoting wound healing and combating bacterial infections in plants.
CC The lytic peptides can also be used for combating protozoal, fungal,
CC viral or bacterial infections or neoplasias in mammals and plants.
CC Lytic peptide-ubiquitin fusion proteins are suitable for production in
CC bacterial hosts. Bacteria lack the hydrolase which cleaves the peptide
CC from ubiquitin, and therefore the active (and cytotoxic) lytic peptide
CC will not be released in the host cells. The recombinantly produced
CC lytic peptide can be retrieved from the fusion protein by cleavage in
CC vitro.
XX
SQ Sequence 21 AA:

Alignment Scores:
Pred. No.: 325 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 17 Gaps: 0

US-09-817-318-1 (1-780) x AAR90756 (1-21)

QY 760 GTAAAGCGCAAAAAAAAAA 780
|||||
Db 15 VallySAAlaLysLysLysLys 21

RESULT 23
AAM39773
ID AAM39773 standard; peptide: 21 AA.
XX
AC AAM39773;
XX
DT 20-MAY-1998 (first entry)
XX
DE Synthetic lysine-rich lytic peptide 24.
XX
KW Lytic peptide; lysine-rich; proteolytic digestion; methylation; trypsin;
KW protection; amphipathic alpha-helix; beta-pleated sheet; treatment;
KW infection; viral; bacterial; yeast; fungal; protozoan; cancer.
XX
OS Synthetic.
XX
PN US5717064-A.
XX
PD 10-FEB-1998.
XX
PE 24-APR-1995; 95US-0427001.
XX
PR 24-APR-1995; 95US-0427001.
PR 18-NOV-1993; 93US-0148889.
XX
PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.
XX
PI Jaynes JM, Julian GR;
XX
WPI: 1998-158370/14.

XX
PT Cytolytic peptide analogues - with methylated lysine residues to
PT increase protease resistance
XX
PS Claim 1; Column 25-26; 21pp; English.
XX
CC Peptides AAM39750-W39785 represent a family of synthetic lytic peptides
CC which are stabilised by having methylated epsilon-amino groups at
CC their lysine residues. Peptides AAM39750-W39771 and AAM39781-W39785 have
CC a amphipathic alpha-helix conformation. These peptides have enhanced
CC resistance to proteolytic digestion by trypsin. Such peptides can be used
CC to treat infections by lysing bacterial, yeast, fungal and protozoan
CC cells or to treat cancers by lysing neoplastic or transformed cells. They
CC can also be used to treat viral infections by lysing enveloped viruses
CC and virus-infected cells.
XX
SQ Sequence 21 AA:

Alignment Scores:
Pred. No.: 325 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 19 Gaps: 0

US-09-817-318-1 (1-780) x AAM39773 (1-21)

QY 760 GTAAAGCGCAAAAAAAAAA 780
|||||
Db 15 VallySAAlaLysLysLysLys 21

RESULT 24
ABG27285
ID ABG27285 standard; Protein: 28 AA.
XX
AC ABG27285;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #27276.
XX
KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
FN WO200175067-A2.
XX
ED 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
FA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PsDB: AAS91472.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 57644; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.

SO Sequence 28 AA;

Alignment Scores:

Pred. No.:	319	Length:	28
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	22	Gaps:	0

US-09-817-318-1 (1-780) x ABG27285 (1-28)

QY 339 TCACGTGCAAAATTCATGAA 319

Db 16 SerLeuSerLysIleHisGlu 22

RESULT 25

AAV68253

ID AAV68253 standard; Peptide: 38 AA.

AC AAV68253;

DT 13-APR-2000 (first entry)

DE Murine class I molecule related protein sequence SEQ ID NO:85.

KW MHC class I; major histocompatibility complex; microglobulin; antigen;

KW immune response; immunisation; AIDS; multiple sclerosis; toxic shock;

KW cancer; lupus erythematosus; snake bite; cytostatic; antiviral;

KW immunomodulatory; dermatological; immunosuppressive; antiinflammatory;

KW neuroprotective.

OS Mus sp.

PN US6011146-A.

PD 04-JAN-2000.

PE 07-JUN-1995; 95US-0481985.

PR 15-NOV-1991; 91US-0792473.

PR 05-DEC-1991; 91US-0801818.

PA (INSP) INST PASTEUR.

PT histocompatibility complex Class I determinant, used for immune system stimulation, e.g. for treating cancer

XX Disclosure: Column 99-100; 88pp; English.

PS The present invention describes a recombinant DNA molecule (I)
 CC containing a sequence (Ia) that encodes an altered MHC (major
 CC histocompatibility complex) Class I determinant (II) comprises a
 CC polypeptide with alpha1, alpha2, alpha3 and beta2-microglobulin
 CC domains, in which alpha1 and beta2 are covalently linked, thorough C-
 CC and N-termini respectively, via a nucleotide spacer sequence encoding a
 CC polypeptide. (II) includes an antigen-binding site and when (II) and
 CC the antigen are associated they are recognized by a mammalian T cell
 CC receptor (TCR). (I) are used to produce (II) which are used to study
 CC functional interactions between the various MHC domains. They can also
 CC be used to modulate (in vivo or in vitro) the immune system by inducing
 CC an effector response (cytotoxicity, antibody synthesis, phagocytosis)
 CC of immune system cells, typically for treating, or immunising against;
 CC cancer, acquired immune deficiency syndrome, lupus erythematosus,
 CC multiple sclerosis, toxic shock and snake bite, but also for selective
 CC destruction of autoreactive cells, diagnostically to assay T cell
 CC receptors and to raise specific antibodies (useful for diagnosis,
 CC therapy, studying MHC-associated cellular processes and for affinity
 CC purification). AAV57558 and AAV68186 to AAV68316 are sequences used in
 CC the exemplification of the present invention.

SO Sequence 38 AA;

Alignment Scores:

Pred. No.:	313	Length:	38
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	21	Gaps:	0

US-09-817-318-1 (1-780) x AAV68253 (1-38)

QY 625 ACTTGTGATTATCCGCGCC 605

Db 19 ThrSerAspSerLeuPro 25

RESULT 26

AAV52907

ID AAV52907 standard; Peptide: 38 AA.

AC AAV52907;

DT 14-FEB-2000 (first entry)

DE Murine class I molecule H-2K-d peptide SEQ ID NO:85.

KW Major histocompatibility complex; MHC class I; MHC class II; antigen;

KW immune response; diagnosis; antibody; immunisation; autoimmune disease;

KW acquired immune deficiency syndrome; AIDS; cytostatic; dermatological;

KW anti-inflammatory; neuroprotective; immunosuppressive; antithyroid;

KW vaccine; lupus erythematosus; multiple sclerosis; thyroiditis;

KW toxic shock; tumour; snakebite.

OS Mus sp.

PN US5976551-A.

PD 02-NOV-1999.

PE 07-JUN-1995; 95US-0484905.

PR 05-DEC-1991; 91US-0801818.

PR 15-NOV-1991; 91US-0792473.

PA (INSP) INST PASTEUR.

PT histocompatibility complex Class I determinant, used for immune system stimulation, e.g. for treating cancer

DR WPI: 2000-037081/03.
XX
XX Composition containing an antigen and altered major histocompatibility
PT Class II determinant, used to immunize against autoimmune diseases,
PT e.g. acquired immune deficiency syndrome -
XX
XX
PS Disclosure: Column 123-124; 96pp; English.
XX
CC The present invention describes a composition capable of eliciting
CC anti-major histocompatibility (MHC) antibodies. The composition
CC comprises an antigen associated with an altered MHC class II determinant
CC (I) comprising alpha1, alpha2, beta1 and beta2 polypeptide domains
CC encoded by a mammalian MHC class II locus covalently linked to form a
CC polypeptide (I) containing beta2, alpha2, alpha1 and beta1 domains in
CC sequence. The resulting Antigen-MHC complex is recognizable by the T cell
CC receptor. The compositions are used for immunisation against, or
CC treatment of, a wide range of autoimmune diseases, e.g. acquired immune
CC deficiency syndrome (AIDS), lupus erythematosus, multiple sclerosis,
CC thyroiditis, toxic shock, tumour and snakebite, depending on the nature
CC of antigen. (I) is also used to analyse functional interactions between
CC the various domains and for targeting lymphocyte receptors. Antibodies
CC against (I) are produced by usual methods of immunisation or cell fusion,
CC and may be humanised by standard methods. These antibodies are useful for
CC diagnosis (detection or purification of MHC gene products), therapy
CC (neutralising MHC on cell surfaces) and in the study of MHC and cellular
CC processes. AA233240 to AA233242 and AA52840 to AA52970 represent
CC sequences used in the exemplification of the present invention.
XX
XX S0 Sequence 38 AA:
S0
Alignment Scores:
Pred. No.: 313 Length: 38
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 21 Gaps: 0
US-09-817-318-1 (1-780) x AA52907 (1-38)
OY 625 ACTTCTGATTATCCCTGCC 605
IIIIIIIIIIIIIIIIIIII
DB 19 ThrSerAspLeuSerLeuPro 25
RESULT 27
AAB58668
ID AAB58668 standard; protein; 38 AA.
XX
XX AAB58668;
XX
XX 13-MAR-2001 (first entry)
XX
XX Murine class I protein #5.
XX
XX Major histocompatibility complex: MHC class I; immune; snake bite;
KM T cell mediated autoimmune disease; AIDS; lupus erythematosus;
KM toxic shock.
XX
XX Mus musculus.
OS
XX
XX US6153408-A.
PN
XX
XX 28-NOV-2000.
PD
XX
XX 09-JAN-1995; 95US-0370476.
PF
XX
XX 15-NOV-1991; 91US-0792473.
PR 07-SEP-1993; 93US-0117575.
PR 05-DEC-1991; 91US-0801818.
PR 07-JUN-1993; 93US-0072787.
XX
XX (INSP) INST PASTEUR.
PA (INRM) INST NAT SANTE & RECH MEDICAL.
PA

XX
XX Abastado J, Kourilsky P, Castrouge A, Ojcius D, Lone Y, Mottez E;
XX
XX WPI: 2001-060089/07.
DR
XX
XX New altered major histocompatibility complex (MHC) class I determinant
PT useful for eliciting an immune response and/or for immunizing against
PT or treating diseases, for example, multiple sclerosis, AIDS, toxic
PT shock or snake bite -
XX
XX
PS Disclosure: Column 19-26; 105pp; English.
XX
XX
CC The present invention relates to a major histocompatibility complex
CC (MHC) class I determinant, which has alpha_1 alpha_2 alpha_3 and
CC beta2-microglobulin polypeptide domains encoded by a mammalian MHC
CC class I locus. The MHC class I determinants are useful for activating
CC the immune system and presenting antigens to the immune system to
CC elicit an antigenic response. The MHC class I determinants are also
CC useful for treating diseases, e.g. T cell mediated autoimmune disease,
CC AIDS, lupus erythematosus, toxic shock or snake bite. The altered MHC
CC class I determinants and compositions containing antigens bound to
CC the determinants are useful in diagnostic applications, e.g. altered
CC determinants may be used to target lymphocyte receptors and the
CC resulting bound determinant can be assayed.
XX
XX S0 Sequence 38 AA:
S0
Alignment Scores:
Pred. No.: 313 Length: 38
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 22 Gaps: 0
US-09-817-318-1 (1-780) x AAB58668 (1-38)
OY 625 ACTTCTGATTATCCCTGCC 605
IIIIIIIIIIIIIIIIIIII
DB 19 ThrSerAspLeuSerLeuPro 25
RESULT 28
AA568254
ID AA568254 standard; peptide; 42 AA.
XX
XX AAY68254;
XX
XX 13-APR-2000 (first entry)
XX
XX Murine class I molecule related protein sequence SEQ ID NO:86.
XX
XX MHC class I; major histocompatibility complex; microglobulin; antigen;
XX immune response; immunisation; AIDS; multiple sclerosis; toxic shock;
KM cancer; lupus erythematosus; snake bite; cytostatic; antiviral;
KM immunomodulatory; dermatological; immunosuppressive; antiinflammatory;
KM neuroprotective.
XX
XX Mus sp.
OS
XX
XX US6011146-A.
PN
XX
XX 04-JAN-2000.
PD
XX
XX 07-JUN-1995; 95US-0481985.
PF
XX
XX 15-NOV-1991; 91US-0792473.
PR 05-DEC-1991; 91US-0801818.
XX
XX (INSP) INST PASTEUR.
PA (INRM) INST NAT SANTE & RECH MEDICALE.
XX
XX Kourilsky P, Mottez E, Abastado J;
PI

DR	WP_02000-125951/11.
XX	
PT	New recombinant DNA encoding covalently linked form of major histocompatibility complex Class I determinant, used for immune system stimulation, e.g. for treating cancer -
PT	
XX	
PS	Disclosure; Column 99-100; 88pp; English.
XX	
CC	The present invention describes a recombinant DNA molecule (I) containing a sequence (Ia) that encodes an altered MHC (major histocompatibility complex) Class I determinant (II) comprises a polypeptide with alpha1, alpha2, alpha3 and beta2-microglobulin domains, in which alpha1 and beta2 are covalently linked, thorough C- and N-terminal respectively, via a nucleotide spacer sequence encoding a polypeptide. (II) includes an antigen-binding site and when (II) and the antigen are associated they are recognized by a mammalian T cell receptor (TCR). (I) are used to produce (II) which are used to study functional interactions between the various MHC domains. They can also be used to modulate (in vivo or in vitro) the immune system by inducing an effector response (cytotoxicity, antibody synthesis, phagocytosis) of immune system cells, typically for treating, or immunising against; cancer, acquired immune deficiency syndrome, lupus erythematosus, multiple sclerosis, toxic shock and snake bite, but also for selective destruction of autoreactive cells, diagnostically to assay T cell receptors and to raise specific antibodies (useful for diagnosis, therapy, studying MHC-associated cellular processes and for affinity purification). AA57558 and AAY68186 to AAY68316 are sequences used in the exemplification of the present invention.
CC	
CC	
CC	
CC	
SO	Sequence 42 AA:
Alignment Scores:	
Pred. No.:	310 Length: 42
Score:	7.00 Matches: 7
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	2.92% Indels: 0
DB:	Gaps: 0
US-09-817-318-1 (1-780) x AAY68254 (1-42)	
Oy	625 ACTCTGATTTATCCCTGCC 605
Db	23 ThrSerAspLeuSerLeuPro 29
RESULT 29	
AAY52908	
ID	AAY52908 standard; peptide; 42 AA.
AC	
XX	AAY52908;
DT	14-FEB-2000 (first entry)
XX	
DE	Murine class I molecule H-2K-q peptide SEQ ID NO:86.
XX	
KW	Major histocompatibility complex; MHC class I; MHC class II; antigen; immune response; diagnosis; antibody; immunisation; autoimmune disease; acquired immune deficiency syndrome; AIDS; cytostatic; dermatological; anti-inflammatory; neuroprotective; immunosuppressive; antithyroid; vaccine; lupus erythematosus; multiple sclerosis; thyroiditis; toxic shock; tumour; snakebite.
XX	
OS	Mus sp.
XX	
PM	US5976551-A.
PD	
XX	02-NOV-1999.
PF	07-JUN-1995; 95US-0484905.
XX	
PR	05-DEC-1991; 91US-0801818.
PR	15-NOV-1991; 91US-0792473.
XX	

PA	(INSP) INST PASTEUR.
PA	(INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX	
PI	Kourilsky P, Mottez E, Abastado J;
XX	
DR	WPI: 2000-037081/03.
XX	
PT	Composition containing an antigen and altered major histocompatibility
PT	Class II determinant, used to immunize against autoimmune diseases,
PT	e.g. acquired immune deficiency syndrome -
XX	
PS	Disclosure; Column 123-124; 96pp; English.
XX	
CC	The present invention describes a composition capable of eliciting
CC	anti-major histocompatibility (MHC) antibodies. The composition
CC	comprises an antigen associated with an altered MHC class II determinant
CC	(1) comprising alpha1, alpha2, beta1 and beta2 polypeptide domains
CC	encoded by a mammalian MHC class II locus covalently linked to form a
CC	polypeptide (1) containing beta2, alpha2, alpha1 and beta1 domains in
CC	sequence. The resulting Antigen-MHC complex is recognizable by the T cell
CC	receptor. The compositions are used for immunisation against, or
CC	treatment of, a wide range of autoimmune diseases, e.g. acquired immune
CC	deficiency syndrome (AIDS), lupus erythematosus, multiple sclerosis,
CC	thyroiditis, toxic shock, tumour and snakebite, depending on the nature
CC	of antigen. (1) is also used to analyse functional interactions between
CC	the various domains and for targeting lymphocyte receptors. Antibodies
CC	against (1) are produced by usual methods of immunisation or cell fusion,
CC	and may be humanised by standard methods. These antibodies are useful for
CC	diagnosis (detection or purification of MHC gene products), therapy
CC	(neutralising MHC on cell surfaces) and in the study of MHC and cellular
CC	processes. AA233240 to AA233242 and AA252840 to AA252870 represent
CC	sequences used in the exemplification of the present invention.
XX	
SO	Sequence 42 AA;
XX	
Alignment Scores:	
Pred. No.:	310
Score:	7.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	2.92%
DB:	21
	Gaps: 0
US-09-817-318-1 (1-780) x AA252908 (1-42)	
QY	625 ACTTCTGATTATCCGCCC 605
Db	23 ThrseraspleuSerleuPro 29
RESULT 30	
AA287287	
ID	AA287287 standard; Protein; 42 AA.
XX	
AC	AA287287;
XX	
DT	07-NOV-2001 (first entry)
XX	
DE	Human immune/haematopoietic antigen SEQ ID NO:14880.
XX	
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KM	cytostatic; gene therapy; vaccine; metastasis.
XX	
OS	Homo sapiens.
XX	
PD	WO200157182-A2.
XX	
PN	09-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US01354.
XX	
RR	31-JAN-2000; 2000US-0179065.
RR	04-FEB-2000; 2000US-0180628.
RR	24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230457.
PR 06-SEP-2000; 2000US-0230458.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0234999.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237039.

PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249254.
PR 17-NOV-2000; 2000US-0249255.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 03-DEC-2000; 2000US-0251030.
PR 03-DEC-2000; 2000US-0251031.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251899.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX
PI
Rosen CA, Barash SC, Ruben SM;
XX
XX
WPI: 2001-483426/52.
DR
DR
N-PSDB; AAK60068.
DR
XX
XX
Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT
PT
useful for preventing, diagnosing and/or treating cancers and
metastasis -
XX
XX
Claim 11; SEO ID NO 14880; 3071pp + Sequence Listing: English.

XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 42 AA;
Alignment Scores:
Pred. No.: 310 Length: 42
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 22 Gaps: 0
US-09-817-318-1 (1-780) x AAM87287 (1-42)
QY 238 TTACAGACCTTACTCTTCAC 258
Db 10 PherthrapsleuThrLeuHis 16
RESULT 31
AAB58669
ID AAB58669 standard; protein; 42 AA.
XX
AC AAB58669;
XX
DT 13-MAR-2001 (first entry)
XX
DE Murine class I protein #6.
XX
KW Major histocompatibility complex; MHC class I; immune; snake bite;
KM T cell mediated autoimmune disease; AIDS; lupus erythematosus;
XX toxic shock.
XX
OS Mus musculus.
XX
PN US6153408-A.
XX
PD 28-NOV-2000.
XX
PF 09-JAN-1995; 95US-0370476.
XX
PR 15-NOV-1991; 91US-0792473.
PR 07-SEP-1993; 93US-0117575.
PR 05-DEC-1991; 91US-0801818.
PR 07-JUN-1993; 93US-0072787.
XX
PA (INSP) INST PASTEUR.
PA (INRM) INST NAT SANTE & RECH MEDICAL.
XX
PI Abastado J, Kourilsky P, Castrouge A, Ojcius D, Lone Y, Mottez E;
XX
DR WPI; 2001-060089/07.
XX
PT New altered major histocompatibility complex (MHC) class I determinant
PT useful for eliciting an immune response and/or for immunizing against
PT or treating diseases, for example, multiple sclerosis, AIDS, toxic
PT shock or snake bite

XX Disclosure; Column 19-26; 105pp; English.
XX
PS The present invention relates to a major histocompatibility complex
CC (MHC) class I determinant, which has alpha_1 alpha_2 alpha_3 and
CC beta2-microglobulin polypeptide domains encoded by a mammalian MHC
CC class I locus. The MHC class I determinants are useful for activating
CC the immune system and presenting antigens to the immune system to
CC elicit an antigenic response. The MHC class I determinants are also
CC useful for treating diseases, e.g. T cell mediated autoimmune disease,
CC AIDS, lupus erythematosus, toxic shock or snake bite. The altered MHC
CC class I determinants and compositions containing antigens bound to
CC the determinants are useful in diagnostic applications, e.g. altered
CC determinants may be used to target lymphocyte receptors and the
CC resulting bound determinant can be assayed.
XX
SQ Sequence 42 AA;
Alignment Scores:
Pred. No.: 310 Length: 42
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 22 Gaps: 0
US-09-817-318-1 (1-780) x AAB58669 (1-42)
QY 625 ACTTCGATTATCCCTGCC 605
Db 23 ThrSerAspleuSerLeuPro 29
RESULT 32
ABG01742
ID ABG01742 standard; protein; 46 AA.
XX
AC ABG01742;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #1733.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM Food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS65929.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20; SEQ ID No 32101; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantifying a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABB00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

XX SQ Sequence 46 AA:

Alignment Scores:

Pred. No.:	309	Length:	46
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.86%	Indels:	0
DB:	22	Gaps:	0

US-09-817-318-1 (1-780) x ABB01742 (1-46)

OY 701 GTTTAATTCTCCGGTTGA 721

Db 17 ValLeuIleLeuValVal 23

RESULT 33

ABB63839

ID AAB63839 standard; Protein: 71 AA.

XX AAB63839.

XX AC

XX DT

26-MAR-2001 (first entry)

XX DE

Human prostate cancer associated antigen protein sequence SEQ ID NO:1201.

XX KM

Human; breast cancer; gastric cancer; prostate cancer; diagnosis;

XX KW

cancer associated antigen; cytostatic; cancer vaccine.

XX OS

Homo sapiens.

XX PN

WO200073801-A2.

XX PD

07-DEC-2000.

XX PF

26-MAY-2000; 2000WO-US14749.

XX PR

28-MAY-1999; 99US-0136526.

XX PT

10-SEP-1999; 99US-0153454.

XX PA

(LUDW-) LUDWIG INST CANCER RES.

XX PI

Obata Y;

XX DR

WPI; 2001-025274/03.

XX PT

Nucleic acids encoding breast, gastric and prostate cancer associated

XX PT

antigen precursors, useful for diagnosing and treating a condition

XX PT

characterized by expression of an abnormal amount of a protein, e.g.

XX cancer -

XX PS

Example 1; Page 741; 799pp; English.

XX CC

AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014

CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP), respectively.
CC AAB62232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer.

XX SQ Sequence 71 AA:

Alignment Scores:

Pred. No.:	300	Length:	71
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	22	Gaps:	0

US-09-817-318-1 (1-780) x AAB63839 (1-71)

OY 649 AACATACTTGTATTGACA 629

Db 2 AsnIleThrLeuLeuThr 8

RESULT 34

AAB63843

ID AAB63843 standard; Protein: 71 AA.

XX AAB63843.

XX AC

XX DT

26-MAR-2001 (first entry)

XX DE

Human prostate cancer associated antigen protein sequence SEQ ID NO:1205.

XX KM

Human; breast cancer; gastric cancer; prostate cancer; diagnosis;

XX KW

cancer associated antigen; cytostatic; cancer vaccine.

XX OS

Homo sapiens.

XX PN

WO200073801-A2.

XX PD

07-DEC-2000.

XX PF

26-MAY-2000; 2000WO-US14749.

XX PR

28-MAY-1999; 99US-0136526.

XX PT

10-SEP-1999; 99US-0153454.

XX PA

(LUDW-) LUDWIG INST CANCER RES.

XX PI

Obata Y;

XX DR

WPI; 2001-025274/03.

XX PT

Nucleic acids encoding breast, gastric and prostate cancer associated

XX PT

antigen precursors, useful for diagnosing and treating a condition

XX PT

characterized by expression of an abnormal amount of a protein, e.g.

XX cancer -

XX PS

Example 1; Page 742; 799pp; English.

XX CC

AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014

XX CC

represent nucleotide sequences encoding human breast, gastric and

XX CC

prostate cancer associated antigen precursors (CAAP), respectively.

XX CC

AAB62232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970

XX CC

represent human breast, gastric and prostate CAAP protein sequence

XX CC

respectively. CAAPs have cytostatic activity and can be used in the

XX CC

production of cancer vaccines. The human CAAP proteins, peptides, nucleic

XX CC

acids or anti-CAAP antibodies are useful for diagnosing and treating a

XX CC

condition characterised by expression of an abnormal amount of a protein,

XX e.g. cancer.

XX Sequence 71 AA;
 Alignment Scores:
 Pred. No.: 300 Length: 71
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.92% Indels: 0
 DB: 22 Gaps: 0

US-09-817-318-1 (1-780) x AAB63843 (1-71)
 QY 649 AACATACCTTGTATTGACA 629
 DB 2 Asn11etHrLeuLeuLeuThr 8

RESULT 35
 AAM07875
 ID AAM07875 standard; Protein: 87 AA.
 XX
 AC AAM07875;
 XX
 DT 22-JUL-1997 (first entry)
 XX
 DE Protein reactive with multiple sclerosis antibodies.
 XX
 KW Mitochondrion; electron; transport; chain; ND4; antigen; protein;
 KM multiple sclerosis; antibody; immunoassay; detection; damage;
 KW spirochaete; surface; preparation; anti-idiotypic; treatment.
 OS Homo sapiens.
 XX
 PN WO9630398-A1.
 XX
 PD 03-OCT-1996.
 XX
 PF 25-MAR-1996; 96WO-AU00166.
 XX
 PR 18-DEC-1995; 95AU-0007205.
 PR 24-MAR-1995; 95AU-0001974.
 XX
 PA (UYMU-) UNITV MURDOCH.
 XX
 PI Carnegie PR, Sanati MH;
 XX
 DR WPI: 1996-505781/50.
 DR N-PSDB: AAT44603.
 XX
 PT DNA encoding a polypeptide reactive with multiple sclerosis-induced
 PT antibodies - used in an immunoassay to detect mitochondrial damage,
 PT and to prepare anti-idiotypic antibodies for therapy.
 XX
 PS Claim 2; Page 39; 61pp; English.
 XX
 CC The present sequence, which is encoded by a nucleotide sequence
 CC having greater than 98% identity with a portion of the
 CC mitochondrial gene encoding the mitochondrial electron transport
 CC chain protein ND4, is a polypeptide immunologically reactive with
 CC multiple sclerosis (MS) induced antibodies (Ab). The polypeptide
 CC can be used in an immunoassay to detect mitochondrial damage, and
 CC Ab reactive with spirochaete surface and mitochondrial proteins.
 CC It can also be used to prepare anti-idiotypic Ab for MS treatment.
 XX
 SO Sequence 87 AA;

Alignment Scores:
 Pred. No.: 295 Length: 87
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.92% Indels: 0
 DB: 17 Gaps: 0

US-09-817-318-1 (1-780) x AAM07875 (1-87)
 QY 649 AACATACCTTGTATTGACA 629
 DB 77 Asn11etHrLeuLeuLeuThr 83

RESULT 36
 AAG03475
 ID AAG03475 standard; Protein: 87 AA.
 XX
 AC AAG03475;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 7556.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KM gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI: 2000-500381/45.
 DR N-PSDB: AAC03481.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 13; SEQ ID 7556; 71pp + CD-ROM; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX
 SO Sequence 87 AA;

Alignment Scores:
 Pred. No.: 295 Length: 87
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.92% Indels: 0
 DB: 21 Gaps: 0

US-09-817-318-1 (1-780) x AAG03475 (1-87)
 QY 673 TTATCACCCTATAATACCA 653
 DB 23 LeuSerProLeuLysTyrPro 29

```
RESULT 37
AA12861
ID AAY12861 standard; Protein: 88 AA.
XX
AC AAY12861;
XX
DT 21-JUN-1999 (first entry)
XX
DE Human 5' EST secreted protein SEQ ID NO:451.
XX
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
XX
OS Homo sapiens.
XX
PN WO9906549-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-IB01231.
XX
PR 01-AUG-1997; 97US-0905279.
XX
PA (GEST ) GENSET.
XX
PI Duclert A, Dumas Milne Edwards J, Lacroix B;
XX
DR WPI: 1999-153779/13.
XX
DR N-PSDB: AAX51639.
XX
PT New nucleic acids encoding human secreted proteins - obtained from
PT cDNA libraries derived from testis, ovary, uterus and spleen tissue
XX
PS Claim 34; Page 489-490; 522pp; English.
XX
CC AAY51459 to AAX51691 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAY12681 to
CC AAY12913, respectively. The proteins given represent the signal peptide
CC and an N-terminal fragment of a secreted protein. The nucleic acid
CC sequences can be used for producing secreted human gene products. They
CC can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell.
XX
SQ Sequence 88 AA:
XX
Alignment Scores:
Pred. No.: 295 Length: 88
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 20 Gaps: 0
US-09-817-318-1 (1-780) x AAY12861 (1-88)
OY 673 TTATACCTATATAATACCA 653
DB 23 LeuSerProIleLysTyrPro 29
```

```
RESULT 38
AA007086
ID AA007086 standard; Protein: 88 AA.
XX
AC AA007086;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 20978.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
XX
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI: 2001-514838/56.
XX
DR N-PSDB: AA187017.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
PS Claim 20; SEQ ID NO 20978; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 88 AA:
XX
Alignment Scores:
Pred. No.: 295 Length: 88
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 22 Gaps: 0
US-09-817-318-1 (1-780) x AA007086 (1-88)
OY 248 TTACTCTGACATTTGACTTG 268
DB 78 LeuLeuPheThrIleAspIleu 84
RESULT 39
ABP03162
ID ABP03162 standard; Protein: 91 AA.
```

XX ABP03162;
 AC XX
 XX 24-JUN-2002 (first entry)
 DT XX
 XX
 DE Human ORFX protein sequence SEQ ID NO:6306.
 XX
 XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.
 XX
 OS Homo sapiens.
 XX
 XX MO200192523-A2.
 PN
 XX
 PD 06-DEC-2001.
 XX
 XX 29-MAY-2001; 2001WO-US10836.
 PF
 XX 30-MAY-2000; 2000US-206132P.
 PR 29-AUG-2000; 2000US-228716P.
 PR
 XX (CURA-) CURAGEN CORP.
 PA
 XX Shinkets RA, Leach MD;
 PI
 XX WPI: 2002-106308/14.
 DR N-PSDB; ABN18914.
 DR
 XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders
 XX
 PS Disclosure; SEQ ID 6306; 1037pp; English.
 XX
 CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 91 AA;

Alignment Scores:
 Pred. No.: 295
 Score: 7.00 Length: 91
 Percent Similarity: 100.00% Matches: 7
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.92% Indels: 0
 DB: 23 Gaps: 0
 US-09-817-318-1 (1-780) x ABP03162 (1-91)
 QY 736 CCAGATCACCATTATACAC 716
 Db 16 ProGUSerrProTYTyrAsn 22
 RESULT 40
 AAG25318
 ID AAG25318 standard; Protein: 94 AA.
 XX
 AC AAG25318;
 XX
 DT 17-OCT-2000 (first entry)
 DT
 DE Zea mays protein fragment SEQ ID NO: 29332.
 XX
 KW Protein identification: signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence; corn.
 XX
 OS Zea mays subsp. mays.
 XX
 PN EP1033405-A2.
 PD
 XX 06-SEP-2000.
 PD
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 PF
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 17-JUN-1999; 9905-0139453.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139817.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 02-JUL-1999; 9905-0142154.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 03-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 06-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 09-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.

PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149729.
PR 23-AUG-1999; 9905-0149920.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 25-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153078.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156569.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159284.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

Alignment Scores:

Pred. No.:	294	Length:	94
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.86%	Indels:	0
DB:	21	Gaps:	0

US-09-817-318-1 (1-780) x AAG25318 (1-94)

QY 418 TCTTCTGTGAGTTCCTTGCC 438
 ||||||||||||||||||
 Db 79 SerSerValSerSerPheAla 85

RESULT 41

ABB80897
 ID ABB80897 standard; Protein: 97 AA.

XX AC ABB80897;

XX DT 08-OCT-2002 (first entry)

XX DE Murine eotaxin polypeptide.

XX Eotaxin: eosinophil; chemotaxis; cytostatic; antiinflammatory; cardiant;
 KW antiallergic; immunosuppressive; antiasthmatic; antiparasitic; histamine;
 KW dermatological; vasotropic; gene therapy; antianaphylactic; mouse.

XX OS Mus musculus.

XX FH Key Location/Qualifiers

FT Peptide 1..23

FT Protein /note= "signal peptide"

FT 24..97

XX /note= "mature protein"

XX PN US6403782-B1.

XX PD 11-JUN-2002.

XX PF 04-AUG-1999; 99US-0366887.

XX PR 22-JUN-1995; 95US-000449P.

XX PR 01-SEP-1995; 95US-0522713.

XX PA (HARD) HARVARD COLLEGE.

XX PA (GEHO) GEN HOSPITAL CORP.

XX PI Luster AD, Leder P, Rothenberg M, Garcia E;

XX DR WPI; 2002-565447/60.

XX DR N-PSDB; ABN6331.

XX PS Claim 1; Fig 3A; 42pp; English.

XX The invention relates to polynucleotides encoding murine, guinea pig or
 CC human eotaxin polypeptides. The eotaxin polynucleotides are useful for
 CC modulating eosinophil chemotaxis, for increasing eosinophil chemotactic
 CC events, and for improving prognosis with tumours in patients. They are
 CC also useful for treating inflammation and tumourigenesis, and for
 CC reducing inflammation and cytotoxic damage caused by eosinophils, for
 CC e.g. during asthmatic reactions, eosinophilic pneumonia and allergic
 CC diseases, inflammatory bowel diseases, atopic dermatitis, urticaria,
 CC vasculitis, parasitic infections and eosinophil cardiac diseases. The
 CC eotaxin polynucleotides are also useful for modulating histamine release
 CC by modulating eotaxin activity or expression during anaphylaxis,
 CC urticaria and allergic reactions. They are useful for detecting and
 CC monitoring eosinophil mediating conditions, and in anti-cancer gene
 CC therapy. The present sequence represents the murine eotaxin.

XX SQ Sequence 97 AA;

Alignment Scores:

Pred. No.: 293

Score: 7.00

Percent Similarity: 100.00%

Length: 97

Matches: 7

Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.92% Indels: 0

DB: 23 Gaps: 0

US-09-817-318-1 (1-780) x ABB80897 (1-97)

QY 640 TTGTTATTGACAGTAACCTTCT 620

|||||||||||||||||

Db 10 LeuLeuLeuThrValThrSer 16

RESULT 42

ABB80898
 ID ABB80898 standard; Protein: 98 AA.

XX AC ABB80898;

XX DT 08-OCT-2002 (first entry)

XX DE Murine eotaxin polypeptide fragment.

XX Eotaxin: eosinophil; chemotaxis; cytostatic; antiinflammatory; cardiant;
 KW antiallergic; immunosuppressive; antiasthmatic; antiparasitic; histamine;
 KW dermatological; vasotropic; gene therapy; antianaphylactic; mouse.

XX OS Mus musculus.

XX PN US6403782-B1.

XX PD 11-JUN-2002.

XX PF 04-AUG-1999; 99US-0366887.

XX PR 22-JUN-1995; 95US-000449P.

XX PR 01-SEP-1995; 95US-0522713.

XX PA (HARD) HARVARD COLLEGE.

XX PA (GEHO) GEN HOSPITAL CORP.

XX PI Luster AD, Leder P, Rothenberg M, Garcia E;

XX DR WPI; 2002-565447/60.

XX PT New DNA encoding murine, guinea pig or human eotaxin polypeptides,
 PT useful for treating inflammation and tumorigenesis and in anticancer
 PT gene therapy -

XX PS Disclosure; Fig 3B; 42pp; English.

XX The invention relates to polynucleotides encoding murine, guinea pig or
 CC human eotaxin polypeptides. The eotaxin polynucleotides are useful for
 CC modulating eosinophil chemotaxis, for increasing eosinophil chemotactic
 CC events, and for improving prognosis with tumours in patients. They are
 CC also useful for treating inflammation and tumourigenesis, and for
 CC reducing inflammation and cytotoxic damage caused by eosinophils, for
 CC e.g. during asthmatic reactions, eosinophilic pneumonia and allergic
 CC diseases, inflammatory bowel diseases, atopic dermatitis, urticaria,
 CC vasculitis, parasitic infections and eosinophil cardiac diseases. The
 CC eotaxin polynucleotides are also useful for modulating histamine release
 CC by modulating eotaxin activity or expression during anaphylaxis,
 CC urticaria and allergic reactions. They are useful for detecting and
 CC monitoring eosinophil mediating conditions, and in anti-cancer gene
 CC therapy. The present sequence represents the murine eotaxin fragment.

XX SQ Sequence 98 AA;

Alignment Scores:

Pred. No.: 293

Score: 7.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2.92%

Length: 98

Matches: 7

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-817-318-1 (1-780) x ABB80898 (1-98)

OY 640 TTGTTATTCAGACTACTCT 620

Db 10 LeuLeuLeuTrValThrSer 16

RESULT 43

AAU59212
ID AAU59212 standard; Protein: 99 AA.

AC AAU59212;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #20108.

KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;

KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;

KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PE 20-APR-2001; 2001WO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI: 2001-616774/71.

DR N-PSDB; AASS9601.

XX Propionibacterium acnes polypeptides and nucleic acids useful for

PT vaccinating against and diagnosing infections, especially useful for

PT treating acne vulgaris -

XX Example 1: SEQ ID NO 20407; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic

CC polypeptides. The proteins and their associated DNA sequences are used in

CC the treatment, prevention and diagnosis of medical conditions caused by

CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,

CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.

CC P. acnes is also involved in infections of bone, joints and the central

CC nervous system, however it is particularly involved in the inflammatory

CC lesions associated with acne vulgaris. A method for detecting the

CC presence or absence of P. acnes in a patient comprises contacting a

CC sample with a binding agent that binds to the proteins of the invention

CC and determining the amount of bound protein in the sample. The

CC polypeptides may be used as antigens in the production of antibodies

CC specific for P. acnes proteins. These antibodies can be used to

CC downregulate expression and activity of P. acnes polypeptides and

CC therefore treat P. acnes infections. The antibodies may also be used as

CC diagnostic agents for determining P. acnes presence, for example, by

CC enzyme linked immunosorbent assay (ELISA).

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2.92%

DB: 22 Gaps: 0

US-09-817-318-1 (1-780) x AAU59212 (1-99)

OY 162 GATTAGTACGACTGCTATTA 142

Db 24 AspleuValLysLeuLeuLeu 30

RESULT 44

AAO07453
ID AAO07453 standard; Protein: 100 AA.

AC AAO07453;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 21345.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PE 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

DR WPI: 2001-514838/56.

DR N-PSDB; AAI87384.

XX Claim 20: SEQ ID NO 21345; 1399pp + Sequence Listing; English.

CC The invention relates to human polynucleotides (AAI79941-AAI93841) and

CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

Alignment Scores:

Pred. No.: 293

Score: 7.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2.92%

Length: 100

Matches: 7

Conservative: 0

Mismatches: 0

Indels: 0

DB: 22 Gaps: 0

US-09-817-318-1 (1-780) x AA007453 (1-100)

QY 649 AACATTAAGTTGTTATGACA 629
 |||||

Db 31 Asn1eThrLeuLeuThr 37

RESULT 45

ABP31587

ID ABP31587 standard; Protein: 104 AA.

AC ABP31587;

XX

DT 08-JUL-2002 (first entry)

XX

DE Human ORF560 protein, SEQ ID NO:1120.

XX

KW Human: ORF; open reading frame; ORFX; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; hematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokine; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KW vasotropic; antiporiatic; antidiabetic; cyostatic; nootropic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antihyroid; antinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.

XX

OS Homo sapiens.

XX

PN WO200190366-A2.

XX

PD 29-NOV-2001.

XX

PF 24-MAY-2001; 2001WO-US17076.

XX

PR 24-MAY-2000; 2000US-206690P.

XX

PA (CURA-) CUBAGEN CORP.

XX

PI Leach MD, Shinkets RA;

XX

DR MPI: 2002-106200/14.

XX

N-PSDB: ABN75613.

XX

PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation

XX

PS Claim 10; Page 546; 2508bp; English.

XX

CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, hematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokine activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and antiinfective activity, and may also be involved in the determination

CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts. In the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.

XX SQ Sequence 104 AA;

Alignment Scores:

Pred. No.:	292	Length:	104
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.86%	Indels:	0
DB:	23	Gaps:	0

US-09-817-318-1 (1-780) x ABP31587 (1-104)

QY 398 ACTCATTTGTTAACTAGTTT 418
 |||||

Db 52 ThrHisLeuLeuThrSerPhe 58

Search completed: February 4, 2003, 07:21:52
 Job time : 50.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_nzp model

Run on: February 4, 2003, 07:17:28 : Search time 15.5 Seconds

(without alignments)
4174.392 Million cell updates/sec

Title: US-09-817-318-1

Perfect score: 245
Sequence: 1 aactactgtctctcttaag.....taaaagcaaaaaaaaaaaaaa 780

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Word size: 1

Total number of hits satisfying chosen parameters: 225614

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+npz model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09817318/runal_04022003_071721_6612/app-query.fasta.1.967
-DB=SwissProt_40 -QFMT=fastan -SUFFIX=resp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DCCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=45 -MODE=LOCAL -OUTFMT=plc
-NOR-ext -HEAPSIZE=500 -MILEN=0 -MAXLEN=2000000000
-USER=US09817318 -CCGN_1_1_19 -funat_04022003_071721_6612 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAPP -LARGEQUERY -NEC_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WALT -TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FCGAPOP=6 -FCGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	3.3	528	1	ASMA4_YEAST
2	8	3.3	607	1	PEPF_MYCE
3	7	2.9	97	1	EOTA_MOUSE
4	7	2.9	100	1	NULM_PHYIN
5	7	2.9	103	1	NIEW_TRISI
6	7	2.9	105	1	THIL_DICDI
7	7	2.9	117	1	V117_HAEDU
8	7	2.9	134	1	LY6F_MOUSE
9	7	2.9	171	1	PA1A_BACSU
10	7	2.9	173	1	RPOE_BACSU
11	7	2.9	174	1	YPUF_BACSU
12	7	2.9	219	1	C1DA_HUMAN
13	7	2.9	219	1	RRP2_SCHHO
14	7	2.9	220	1	Y069_GAEBL
15	7	2.9	245	1	PHOS_BOVIN
16	7	2.9	245	1	PHOS_CANFA
17	7	2.9	283	1	CHER_BORBU
18	7	2.9	283	1	PSTA_RHILU

19	7	2.9	312	1	KRAW_BUCAP	085295 buchnera ap
20	7	2.9	314	1	REBN_SALTY	P26403 salmonella
21	7	2.9	328	1	HA1Q_MOUSE	P14428 mus musculus
22	7	2.9	332	1	HRDD_STRVF	O60012 streptomyc
23	7	2.9	333	1	MDHC_HUMAN	P40925 homo sapien
24	7	2.9	333	1	MDHC_MOUSE	P14152 mus musculus
25	7	2.9	355	1	C3X1_HUMAN	P49238 homo sapien
26	7	2.9	368	1	HA1D_HUMAN	P01902 mus musculus
27	7	2.9	368	1	HA1W_MOUSE	P03391 mus musculus
28	7	2.9	369	1	HA1B_MOUSE	P01901 mus musculus
29	7	2.9	369	1	HA1K_MOUSE	P04223 mus musculus
30	7	2.9	371	1	CYB_CANAS	O48023 candoia asp
31	7	2.9	385	1	O46A_DROME	P81919 drosophila
32	7	2.9	411	1	Y360_MYCE	O49426 mycoplasma
33	7	2.9	415	1	AMS0_ERYAM	O46637 erwinia amy
34	7	2.9	435	1	POTE_HAETN	P44766 haemophilus
35	7	2.9	447	1	G6P1_LACIA	P81181 lactococcus
36	7	2.9	453	1	YKK8_YEAST	P34252 saccharomyc
37	7	2.9	459	1	N04M_HUMAN	P03905 homo sapien
38	7	2.9	460	1	N04M_BRARE	O9miy1 brachydanio
39	7	2.9	461	1	PRTC_RAT	P31394 rattus norv
40	7	2.9	463	1	FMNL_HUMAN	O93466 homo sapien
41	7	2.9	475	1	AP54_YEAST	O00776 saccharomyc
42	7	2.9	486	1	YDE5_SCHHO	O10439 schizosacch
43	7	2.9	513	1	VGA_BPPHX	P03631 bacterioph
44	7	2.9	519	1	SC59_YEAST	P20048 saccharomyc
45	7	2.9	522	1	VGA_BPS13	P07928 bacterioph

ALIGNMENTS

RESULT 1

ASMA4_YEAST
ID ASMA4_YEAST STANDARD: PRT: 528 AA.

AC Q05166: Q12456;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE ASMA4 protein.

GN ASMA4 OR YDL088C OR D2420.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

XX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=95166178; PubMed=7862092;

RA Giot L., Simon M., Dubois C., Faye G.;

*Suppressors of thermosensitive mutations in the DNA polymerase delta

gene of Saccharomyces cerevisiae.*;

Mol. Gen. Genet. 246:212-222(1995).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN=S288c / FY1679;

RA Boskovic J., Saiz J.E., Soler-Mira A., Garcia-Cantalejo J.,

Revuelta J.L., Jimenez A., Ballesta J.P.G., del Rey F., Remacha M.;

Submitted (FEB-1996) to the EMBL/Genbank/DBJ databases.

RL [3]

RP SEQUENCE FROM N.A.

W submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.

-I- MISCELLANEOUS: SUPPRESSOR OF THERMOSENSITIVE MUTATIONS IN THE DNA

POLYMERASE DELTA GENE (POL3).

-I- SIMILARITY: TO YEAST YMR153W.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: X76709; CAA54130.1; -
DR EMBL: X95644; CAA64923.1; -
DR EMBL: 274136; CAA98654.1; -
DR SGD: S0002246; ASW4.
FT DOMAIN 38 46 POLY-GLN.
FT 84 87 POLY-ASN.
FT 114 114 F-> L (IN REF. 1).
FT CONFLICT 446 528 PAGHAGNPTNISPPIVANSPNKRLVDIGKLPFMONAGPNS
FT CONFLICT 446 528 NIPNLRLNLESKMRQOEAKRYRNPAGTGHKLSNMLPFGMND
FT L-> LPMVLVIOQFQVQ (IN REF. 1).
SQ SEQUENCE 528 AA: 58793 MW: 84FE1F07FC1B0173 CRC64;

Alignment Scores:

Pred. No.:	6.98	Length:	528
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.27%	Indels:	0
DB:	1	Gaps:	0

US-09-817-318-1 (1-780) x ASW4_YEAST (1-528)

QY 670 ATAATTCATTACGACATTC 693

Db 170 IleaSerLeuSerAspHe 177

RESULT 2

PEPF_MYCGE	STANDARD:	PRT:	607 AA.
ID P47429:			

AC 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Oligendopeptidase F homolog (EC 3.4.24.-).

GN PEPF OR MG183.

OS Mycoplasma genitalium.

OC Bacteria: Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI_TaxID=2097;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 33530 / G-37;

RX MEDLINE=96026346; PubMed=7569993;

RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,

RA Fleischmann J.L., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,

RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,

RA Nguyen D.T., Uterback T.R., Sauder D.M., Phillips C.A., Merrick J.M.,

RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Luster T.S.,

RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;

RT "The minimal gene complement of Mycoplasma genitalium.";

RL Science 270:397-403(1995).

RN [2]

RP SEQUENCE OF 10-111 FROM N.A.

RC STRAIN=ATCC 33530 / G-37;

DR InterPro: IPR004438; PEPF.
DR InterPro: IPR001567; Peptidase_M3.
DR InterPro: IPR000130; Zn_Mpeptidse.
DR Pfam: PF01432; Peptidase_M3; 1.
DR TIGRfams: TIGR00181; PEPF; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Hydrolyase; Metalloprotease; Zinc; Complete proteome.
FT METAL 384 384 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 385 385 BY SIMILARITY.
FT METAL 388 388 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 391 391 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 607 AA: 70869 MW: 47IDFF73A4ED7 CRC64;

Alignment Scores:

Pred. No.:	6.8	Length:	607
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.33%	Indels:	0
DB:	1	Gaps:	0

US-09-817-318-1 (1-780) x PEPF_MYCGE (1-607)

QY 150 TTGCTATATGCTATTATGACTA 127

Db 419 LeuLeuLeuCysTyrTyrGluLeu 426

RESULT 3

EOTA_MOUSE	STANDARD:	PRT:	97 AA.
ID EOTA_MOUSE			

AC P48298:

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Eotaxin precursor (Small inducible cytokine A11) (CCL11) (Eosinophil

DE chemotactic protein).

GN SCY11.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Lung;

RX MEDLINE=96004658; PubMed=7568052;

RA Rothenberg M.E., Luster A.D., Leder P.;

RA "Murine eotaxin: an eosinophil chemoattractant inducible in

RT endothelial cells and in interleukin 4-induced tumor suppression.";

RL Proc. Natl. Acad. Sci. U.S.A. 92:8960-8964(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Lung;

DR InterPro: IPR001567; Peptidase_M3.
DR InterPro: IPR000130; Zn_Mpeptidse.
DR Pfam: PF01432; Peptidase_M3; 1.
DR TIGRfams: TIGR00181; PEPF; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Hydrolyase; Metalloprotease; Zinc; Complete proteome.
FT METAL 384 384 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 385 385 BY SIMILARITY.
FT METAL 388 388 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 391 391 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 607 AA: 70869 MW: 47IDFF73A4ED7 CRC64;

Alignment Scores:

Pred. No.:	6.8	Length:	607
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.33%	Indels:	0
DB:	1	Gaps:	0

US-09-817-318-1 (1-780) x PEPF_MYCGE (1-607)

QY 150 TTGCTATATGCTATTATGACTA 127

Db 419 LeuLeuLeuCysTyrTyrGluLeu 426

RESULT 3

EOTA_MOUSE	STANDARD:	PRT:	97 AA.
ID EOTA_MOUSE			

AC P48298:

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Eotaxin precursor (Small inducible cytokine A11) (CCL11) (Eosinophil

DE chemotactic protein).

GN SCY11.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Lung;

RX MEDLINE=96158746; PubMed=8574847;

RA Gonzalez J.A., Jia G.-O., Aguirre V., Friend D., Coyle A.J.,

RA Jenkins N.A., Lin G.-S., Katz H., Lichtman A., Copeland N.G., Kopf M.,

RA "Mouse eotaxin expression parallels eosinophil accumulation during

RT lung allergic inflammation but it is not restricted to a Th2-type

RL response.";

RN [2]

RP SEQUENCE OF 10-111 FROM N.A.

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U26426; AAC52256.1; -
 DR EMBL: U40672; AAA99776.1; -
 DR HSSP: P51671; 1E0T.
 DR MGD: MG1:103576; Scyall.
 DR InterPro: IPR000827; CC_Chemkine_sml.
 DR InterPro: IPR001811; Chemokine_IL8.
 DR Pfam: PF00048; IL8; 1.
 DR SMART: SM00199; SCY; 1.
 DR PROSITE: PS00472; SMALL_CYTOKINES_CC; 1.
 KW Eosinophil: Cytokine; Chemotaxis; Glycoprotein; Signal;
 KW Inflammatory response.
 FT SIGNAL 1 23
 FT CHAIN 24 97
 FT DISULFID 32 57
 FT DISULFID 33 73
 FT CARBOHYD 94 94
 SQ SEQUENCE 97 AA: 10893 MW: 36C9812107FC6CA7 CRC64; (BY SIMILARITY).

Alignment Scores:
 Pred. No.: 103 Length: 97
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.92% Indels: 0
 DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x EOTL_MOUSE (1-97)

QY 640 TTGTTATTGACAGTACTCT 620
 ||||||||||||||||
 Db 10 LeuLeuLeuThrValThrSer 16

RESULT 4
 NULM_PHYIN STANDARD; PRT; 100 AA.
 ID NULM_PHYIN
 AC Q37598;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
 GN NAD4L OR NAD4L.
 OS Phytophthora infestans (potato late blight fungus).
 OC Mitochondrion.
 OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
 OC Phytophthora.
 OC NCBI_TaxID=4787;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 16981 / West Virginia 4;
 RA Lang B.F., Forget L.;
 RT "The mitochondrial genome of *Phytophthora infestans*."
 RL (in) O'Brien S.J. (eds.);
 RL Genetic maps, pp.3:133-135, Cold Spring Harbor Laboratory Press,
 RL New-York (1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 16981 / West Virginia 4;
 RA Chesnick J.M., Tuxbury K., Coleman A., Burger G., Lang B.F.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U17009; AAF24790.1; -
 DR EMBL: U54634; AAB00440.1; -
 DR InterPro: IPR003215; NADH_dh_ubiq1.
 DR InterPro: IPR001133; Oxidored4L.
 DR InterPro: IPR003214; Oxidored4L.
 DR Pfam: PF00420; Oxidored_q2; 1.
 DR ProDom: PD002107; NADH_dh_ubiq1; 1.
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
 SQ SEQUENCE 100 AA: 11277 MW: 199092C0874BCB15 CRC64;

Alignment Scores:
 Pred. No.: 103 Length: 100
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.92% Indels: 0
 DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x NULM_PHYIN (1-100)

QY 114 TATAAATAAAGTACTCTCA 94
 ||||||||||||||||
 Db 93 TyrIlySileAsnLeuLeuSer 99

RESULT 5
 NIFW_TRIS1 STANDARD; PRT; 103 AA.
 ID NIFW_TRIS1
 AC Q9KJL4;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Nitrogenase stabilizing/protective protein nifw.
 GN NIFW.
 OS Trichodesmium sp. (strain IMS101).
 OC Bacteria; Cyanobacteria; Oscillatoriales; Trichodesmium.
 OC NCBI_TaxID=57878;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dominic B., Zani S., Chen Y.-B., Mellon M.T., Zehr J.P.;
 RT "Organization of the nif genes of the nonheterocystous cyanobacterium
 RT Trichodesmium sp. IMS101."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: May protect the nitrogenase Fe-mo protein from oxidative
 CC damage (By similarity).
 CC -!- SUBUNIT: Homotrimer; associates with nifd (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE NIFW FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF167538; AAF82643.1; -
 DR InterPro: IPR004893; Nifw.
 DR Pfam: PF03206; Nifw; 1.
 KW Nitrogen fixation.
 SQ SEQUENCE 103 AA: 11819 MW: BE423C1E4309D2A6 CRC64;

Alignment Scores:
 Pred. No.: 102 Length: 103
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.92% Indels: 0
 DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x NIFW_TRIS1 (1-103)

OY 542 ACAGTCATGTAATGATTG 522
 DB 28 ThrValAsnValAsnArgLeu 34

RESULT 6
 TH1L_DICDI STANDARD; PRT; 105 AA.
 ID TH1L_DICDI
 AC P29445;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Thioredoxin 1.
 GN TRXA OR TRX1.
 OS Dictyostellium discoideum (slime mold).
 OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92250653; PubMed=1577820;
 RA Wettelauer B., Jacquot J.-P., Veron M.;
 RT "Thioredoxins from Dictyostellium discoideum are a developmentally
 regulated multigene family";
 RL J. Biol. Chem. 267:9895-9904(1992).
 CC -!- FUNCTION: Participates in various redox reactions through the
 reversible oxidation of its active center dithiol to a disulfide
 and catalyses dithiol-disulfide exchange reactions.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: M91384; AAA3258.1; -.
 DR PIR: A46264; A46264.
 DR HSSP: P10599; 1ERV.
 DR DICTYDB: D05029; trxa.
 DR InterPro: IPR000063; Thiored.
 DR Pfam: PF00085; Thiored.1.
 DR PRINTS: PR00421; THIOREDIXIN.
 DR PROSITE: PS00194; THIOREDIXIN, 1.
 DR Redox-active center; Electron transport; Multigene family.
 FT DISULFID 32 35 REDOX-ACTIVE (BY SIMILARITY).
 SQ SEQUENCE 105 AA; 11926 MW; 0BC12C2867CEB1F5 CRC64;

Alignment Scores:
 Pred. No.: 102 Length: 105
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.92% Indels: 0
 DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x TH1L_DICDI (1-105)
 OY 686 CTCAGTAATGATTATTCACC 666
 DB 45 LeuSerAsnGluPheIleThr 51

RESULT 7
 Y117_HAEDU STANDARD; PRT; 117 AA.
 ID Y117_HAEDU
 AC O30825;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Hypothetical protein HYPO117.
 OS Haemophilus ducreyi.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.

OX NCBI_TaxID=730;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=35000;
 RA San Mateo L.R., Toffer K.L., Kawula T.H.;
 RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE UPF0076 (UK114) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: AF017750; AAC46218.1; -.
 DR InterPro: IPR000543; Y3GF-like.
 DR Pfam: PF01042; UPF0076; 1.
 DR PROSITE: PS01094; UPF0076; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 117 AA; 13252 MW; BB3FB8C4827940B2 CRC64;

Alignment Scores:
 Pred. No.: 99.8 Length: 117
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.92% Indels: 0
 DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x Y117_HAEDU (1-117)
 OY 163 ACATTAGTGAAGTGTCTATT 143
 DB 10 ArgPheSerGluValAlaIle 16

RESULT 8
 LY6F_MOUSE STANDARD; PRT; 134 AA.
 ID LY6F_MOUSE
 AC P35460;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lymphocyte antigen Ly-6F.1 precursor.
 GN LY6F.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Liver.
 RX MEDLINE=93294293; PubMed=8515066;
 RA Fleming T.J., O'Huigin C., Malek T.R.;
 RT "Characterization of two novel Ly-6 genes. Protein sequence and
 potential structural similarity to alpha-bungarotoxin and other
 neurotoxins";
 RT J. Immunol. 150:5379-5390(1993).
 RL -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
 similarity).
 CC -!- SIMILARITY: CONTAINS 1 UPAR/LY6 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: X70922; CAA50270.1; -.
 DR EMBL: X70918; CAA50270.1; JOINED.


```

DR EMBL: X70919: CAAS0270.1; JOINED.
DR MGD; MGI:109441; LY6f.
DR InterPro: IPR001526; LY6_UPAR.
DR Pfam: PF00021; UPAR_LY6; 1.
DR SMART: SM00134; LY; 1.
DR PROSITE: PS00983; LY6_UPAR; 1.
KW T-cell; Signal; Antigen; Multigene family; GPI-anchor.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 219 LYMPHOCYTE ANTIGEN LY-6F.1.
FT PROPEP 2120 134 REMOVED IN MATURE FORM (POTENTIAL).
FT DOMAIN 27 119 UPAR/LY6.
FT DISULFID 29 53 BY SIMILARITY.
FT DISULFID 32 41 BY SIMILARITY.
FT DISULFID 46 74 BY SIMILARITY.
FT DISULFID 78 98 BY SIMILARITY.
FT DISULFID 99 104 BY SIMILARITY.
FT LIPID 119 119 GPI-ANCHOR (POTENTIAL).
SO SEQUENCE 134 AA; 14599 MW; 3FF9DD32FA9183CA CRC64;

Alignment Scores:
Pred. No.: 97.3 Length: 134
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x LY6F_MOUSE (1-134)

QY 701 GTTTAATTCCTCGTGTGA 721
|||||
Db 11 ValleuleleuValVal 17

RESULT 9
PAIA_BACSU STANDARD; PRT; 171 AA.
ID PAIA_BACSU STANDARD; PRT; 171 AA.
AC P1340; O32112;
DT 01-MAY-1991 (Rel. 18, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protease synthase and sporulation negative regulatory protein PAI 1.
DE PAIA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.
RC STRAIN=168 / DB104;
RX MEDLINE=90202692; PubMed=2108124;
RA Honjo M., Nakayama A., Fukazawa K., Kawamura K., Ando K., Hori M., Furutani Y.;
RT "A novel Bacillus subtilis gene involved in negative control of sporulation and degradative-enzyme production.";
RL J. Bacteriol. 172:1783-1790(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Aaevedo V., Bertoletti M.G., Bessieres P., Bolotin A., Borchert S., Borst R., Boursier L., Brans A., Braun M., Britnell S.C., Bron S., Brouillet S., Brusch C.V., Caldwell B., Capuano N.J., Carter N.M., Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoft A., Ehlich S.D., Emerson P.T., Entian K.D., Errington J., Fabre C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Kletter-Blanchard M., Klein C., Kobayashi Y., Koeltter P., Koningsstein G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,

```

```

RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetle D., Potwolk S., Prescott A.M., Prescan E., Pujić P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadde Y., Sato T., Scanlan E., Schleich S., Schreier R., Scoffone F., Sekiguchi J., Sekowska A., Serró S.J., Serró P., Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpsita P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarelli A., Viari A., Wambutt R., Wedler H., Weitzengraber T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.";
RL Nature 390:249-256(1997).
CC -1- FUNCTION: INVOLVED IN NEGATIVE CONTROL OF SPORULATION AND DEGRADATIVE-ENZYME PRODUCTION.
CC -1- SIMILARITY: STRONG. TO L.DELBRUECKII HYPOTHETICAL PROTEIN IN PEPI 3'REGION (AC P46543). SOME, TO M.TUBERCULOSIS RV2669.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M36471; AAA22638.1; -.
DR EMBL: 299120; CAB15205.1; -.
DR PIR: A35145; A35145.
DR Subtilist; BG10695; paia.
DR InterPro: IPR000182; GCN5acetyltransf.
DR Pfam: PF00583; Acetyltransf; 1.
KW DNA-binding; Transcription regulation; Repressor; Sporulation;
KM Complete proteome.
FT INIT MET 0
FT DNA_BIND 97 121 H-T-H MOTIF.
FT CONFLICT 26 26 T->I (IN REF. 1).
SO SEQUENCE 171 AA; 19883 MW; DB5EB68E9647C9D CRC64;

Alignment Scores:
Pred. No.: 93.1 Length: 171
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x PAIA_BACSU (1-171)

QY 504 TCTTCCAAAGCATGGCTT 484
|||||
Db 100 SerpHeGlnYSHISgLYeu 106

RESULT 10
RPOE_BACSU STANDARD; PRT; 173 AA.
ID RPOE_BACSU STANDARD; PRT; 173 AA.
AC P12464;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA-directed RNA polymerase delta subunit (RNAP delta factor).
DE RPOE.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88329737; PubMed=2843435;
RA Lampe M., Blinle C., Schmidt R., Losick R.;
RT "Cloned gene encoding the delta subunit of Bacillus subtilis RNA

```

```

RT polymerase."
RL Gene 67:13-19(1988).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-168:
RX MEDLINE=98015417; PubMed=9353933;
RA Prescan E., Moszer I., Bourcier L., Cruz Ramos H.C., De la Fuente V.,
RA Hullo M.-F., Lelong C., Schleich S., Sekowska A., Song B.H.,
RA Villani G., Kunst F., Danchin A., Glaser P.;
RT "The Bacillus subtilis genome from gerBC (311 degrees) to l1cR (334
RT degrees)."
RL Microbiology 143:3313-3328(1997).
RN [3]
RN SEQUENCE OF 143-173 FROM N.A.
RC STRAIN-168 / JH642;
RX MEDLINE=98314920; PubMed=2457578;
RA Treich K., Chapman J.W., Pigot P., Lecocq D., Hoch J.A.;
RT "Complete sequence and transcriptional analysis of the spo0F region
RT of the Bacillus subtilis chromosome."
RL J. Bacteriol. 170:4194-4208(1988).
RN [4]
RN CHARACTERIZATION
RX MEDLINE=95404613; PubMed=7545758;
RA Lopez de Saro F.J., Woody A.V., Helmann J.D.;
RT "Structural analysis of the Bacillus subtilis delta factor: a protein
RT polyanion which displaces RNA from RNA polymerase."
RL J. Mol. Biol. 252:189-202(1995).
RN [5]
RN CHARACTERIZATION
RX MEDLINE=99269144; PubMed=10336502;
RA Lopez de Saro F.J., Yoshikawa N., Helmann J.D.;
RT "Expression, abundance, and RNA polymerase binding properties of the
RT delta factor of Bacillus subtilis."
RL J. Biol. Chem. 274:15953-15958(1999).
RN [6]
RN FUNCTION: Participates in both the initiation and recycling phases
RN of transcription. In the presence of the delta subunit, RNAP
RN displays an increased specificity of transcription, a decreased
RN affinity for nucleic acids, and an increased efficiency of RNA
RN synthesis because of enhanced recycling. May function in sigma
RN factor switching. It displaces RNA bound to RNA polymerase in a
RN binary complex.
CC -1- SUBUNIT: RNAP is composed of a core of 2 alpha, a beta and a beta'
CC subunits. The core is associated with a delta subunit and one of
CC several sigma factors.
CC -1- SIMILARITY: BELONGS TO THE RPOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M21677; AAA22710.1; -
DR EMBL: 249782; CAA89869.1; -
DR EMBL: M22039; AAA16800.1; -
DR EMBL: 299123; CAB15744.1; -
DR PIR: JTO302; JTO302.
DR PIR: H32354; H32354.
DR Subtilisin; BG10409; pPOE.
KW DNA-directed RNA polymerase; Transcription; Complete proteome.
FT DOMAIN 107 173 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 173 AA; 20399 MW; 0F5818802654417D CRC64;

```

```

Alignment Scores:
Pred. No.: 92.9 Length: 173
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 1 Gaps: 0

```

```

US-09-817-318-1 (1-780) x RPOE_BACSU (1-173)
QY 760 GTAAGCGCAAAAAAAAAA 780
DB 95 VALLYSALATSLYSLYSLYS 101
RESULT 11
YPUF_BACSU STANDARD: PRT; 174 AA.
AC P17617;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein ypuF (ORF6).
GN ypuF.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=1423;
RN [1]
RN SEQUENCE FROM N.A.
RX STRAIN-168 / Marburg;
RC MEDLINE=95020538; PubMed=7934829;
RA Sorokin A.V., Zumbstein E., Azevedo V., Ehrlich S.D., Serror P.;
RT "The organization of the Bacillus subtilis 168 chromosome region
RT between the spoVA and serA genetic loci, based on sequence data."
RL Mol. Microbiol. 10:385-395(1993).
RN [2]
RN SEQUENCE FROM N.A.
RX STRAIN-168 / SHGW;
RC Mironov V.N.;
RT Thesis (1989), USSR Academy of Sciences, Russia.
RN [3]
RN SEQUENCE FROM N.A.
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borris R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Ertan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,
RA Gilm S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Bianchard M., Klein C.,
RA Kobayashi Y., Koelter P., Konigstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Muel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Portwilk S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Koche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiyuchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Takakoshi A., Tanaka T., Tempista P., Tognoni A.,
RA Toseco V., Uchiyama S., Vandenbol M., Vannier F., Vasseroiti A.,
RA Viari A., Wambuit R., Wedler E., Wedler H., Weitzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to licensee@lsb.ch).
CC -----
DR EMBL: L092228; AAA67486.1; -
DR DR EMBL: X51510; CAA35883.1; -
DR EMBL: Z99116; CAB14255.1; -
DR PIR: S45548; S45548.
DR Subtilist: BG10523; YPUF.
KM Hypothetical protein: Complete proteome.
SQ SEQUENCE 174 AA; 21021 MW; 2D969CFAFCB7EBAE CRC64;

Alignment Scores:
Pred. NO.: 92.8 Length: 174
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x YPUF_BACSU (1-174)
QY 603 TTATCTATATCTTTTGATT 583
Db 141 LeuSerAsnIlePheIleuIle 147

RESULT 12
CIDA_HUMAN STANDARD; PRT; 219 AA.
ID 060543;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cell death activator CIDE-A (Cell death-inducing DFFA-like effector
A).
GN CIDEA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=98232498; PubMed=9564035;
RX Inohara N., Koseki T., Chen S., Wu X., Nunez G.;
RT "CIDE, a novel family of cell death activators with homology to the 45
RT KDA subunit of the DNA fragmentation factor.";
RL EMO J. 17:2526-2533(1998).
CC -1- FUNCTION: ACTIVATES APOPTOSIS.
CC -1- SUBUNIT: INHIBITED BY DFFB.
CC -1- SIMILARITY: CONTAINS 1 CIDE-N DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licensee@lsb.ch).
CC -----
DR EMBL: AF041378; AAC34987.1; -
DR HSSP: Q90HD4; ID4B.
DR Genew: HGNC:1976; CIDEA.
DR MIM: 604440; -
DR InterPro: IPR003508; CAD.
DR Pfam: PF02017; CIDE-N; 1.
DR SMART: SM00266; CAD; 1.
KM Apoptosis.
FT DOMAIN 33 110 CIDE-N.
SQ SEQUENCE 219 AA; 24686 MW; 05F704823CE71C0E CRC64;

Alignment Scores:
Pred. NO.: 88.9 Length: 219
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

```

```

Query Match: 2.86% Indels: 0
DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x CIDA_HUMAN (1-219)
QY 281 ACAAGAGGCTCTGTTTACA 301
Db 22 ThrIysArgValIleuPheThr 28

RESULT 13
REP2_SCHPO STANDARD; PRT; 219 AA.
ID 009824;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transcriptional activator protein rep2.
GN REP2 OR SPBC2F12.11C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96030785; PubMed=7588609;
RA Nakashima N., Tanaka K., Sturm S., Okayama H.;
RT "Fission yeast Rep2 is a putative transcriptional activator subunit
RT for the cell cycle 'start' function of Res2-Cdc10.";
RL EMO J. 14:4794-4802(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rotherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skellon J., Simmonds M., Squares R., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voiclaert G., Aert R., Robben J., Grymoprez B.,
RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wandt R., Purnelle B.,
RA Gofeu A., Cadieu E., Dreano S., Gloux S., Delaure V., Motlier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Bento J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR WHICH INTERACTS WITH THE MCB
CC BINDING SUBUNIT COMPLEX FORMED BY RES2 AND CDC10. REP2 IS REQUIRED
CC FOR THE MITOTIC CELL CYCLE START.
CC -1- SIMILARITY: TO REP1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licensee@lsb.ch).
CC -----

```

DR EMBL: X91044; CAA62504.1; -
 DR EMBL: Z97211; CAB10158.1; -
 DR TRANSFAC: T04219; -
 DR InterPro: IPR000822; znf_C2H2.
 DR Pfam: PF00096; zf-C2H2.1.
 KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
 KW Phosphorylation; Mitosis; Cell cycle.
 FT ZN_FING 177 POTENTIAL.
 SQ SEQUENCE 219 AA; 24668 MW; 77F1IA06E722656 CRC64;

Alignment Scores:

Pred. No.:	88.9	Length:	219
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.86%	Indels:	0
DB:	1	Gaps:	0

US-09-817-318-1 (1-780) x REP2_SCHPO (1-219)

OY 418 TCTTGTGAGTCTTGGC 438

Db 42 SetServerSerPheala 48

RESULT 14

ID Y069_CAEEL STANDARD; PRT; 220 AA.

AC P34608;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical 23.9 kDa protein ZK1098.9 in chromosome III.

GN ZK1098.9.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Bristol N2; PubMed-7906398;

RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,

RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,

RA Craton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,

RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,

RA Johnston L., Jones A., Kershaw J., Kirsten J., Laister N.,

RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,

RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,

RA Sims M., Smailton N., Smith A., Smith M., Sonnenhammer E., Staden R.,

RA Sulston J., Thierry-Mieg J., Thomas K., Vaughan K.,

RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,

RA Wohldman P.;

RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

RL Nature 368:32-38(1994).

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.86% Indels: 0
 DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x Y069_CAEEL (1-220)

OY 546 TGCATTGATGCTAGATGTT 566

Db 183 Cysillevalleuanspyal 189

RESULT 15

ID PHOS_BOVIN STANDARD; PRT; 245 AA.

AC P19632; P20940; Q28160;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-MAY-1991 (Rel. 18, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Phosducin (PHD) (33 kDa phototransducing protein) (MEKA protein).

GN PDC.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

RN [1]

RP SEQUENCE AND SEQUENCE OF 228-245 FROM N.A.

RC TISSUE-Retina;

RA MEDLINE-90368806; PubMed-2203790;

RA Lee R.H., Fowler A., McGinnis J.F., Lolley R.N., Craft C.M.;

RT "Amino acid and cDNA sequence of bovine phosducin, a soluble

phosphoprotein from photoreceptor cells."

RL J. Biol. Chem. 265:15867-15873(1990).

RN [2]

RP SEQUENCE FROM N.A.

RC MEDLINE-89364083; PubMed-2770450;

RA Kuo C.-H., Akiyama M., Miki N.;

RT "Isolation of a novel retina-specific clone (MEKA cDNA) encoding a

photoreceptor soluble protein."

RL Brain Res. Mol. Brain Res. 6:1-10(1989).

RN [3]

RP SEQUENCE OF 8-245 FROM N.A.

RC TISSUE-Retina; and Pineal gland;

RA MEDLINE-91007277; PubMed-2210381;

RA Abe T., Nakabayashi H., Tamada H., Takagi T., Sakuragi S., Yamaki K.,

RA Shiohara T.;

RT "Analysis of the human, bovine and rat 33-kDa proteins and cDNA in

retina and pineal gland."

RL Gene 91:209-215(1990).

RN [4]

RP PHOSPHORYLATION OF SER-73.

RC MEDLINE-90368805; PubMed-2394752;

RA Lee R.H., Brown B.M., Lolley R.N.;

RT "Protein kinase A phosphorylates retinal phosducin on serine 73 in

situ."

RL J. Biol. Chem. 265:15860-15866(1990).

RN [5]

RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF COMPLEX WITH G-BETA AND

G-GAMMA.

RA MEDLINE-98416696; PubMed-9739091;

RA Loew A., Ho Y.K., Blundell T., Bax B.;

RT "Phosducin induces a structural change in transducin beta gamma."

RL Structure 6:1007-1019(1998).

CC -! FUNCTION: MAY PARTICIPATE IN THE REGULATION OF VISUAL

PHOTOTRANSDUCTION OR IN THE INTEGRATION OF PHOTORECEPTOR

METABOLISM.

CC -! SUBUNIT: Forms a complex with the beta and gamma subunits of

the GTP-binding protein, transducin.

CC -! SUBCELLULAR LOCATION: OUTER AND INNER SEGMENTS OF THE ROD CELLS.

CC -! MISCELLANEOUS: LIGHT-INDUCED CHANGES IN CYCLIC NUCLEOTIDE LEVELS

MODULATE THE PHOSPHORYLATION OF THIS PROTEIN BY CAMP KINASE.

CC -! SIMILARITY: BELONGS TO THE PHOSDUCIN FAMILY.

CC -! CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN IN THE C-TERMINUS

Alignment Scores:

Pred. No.:	88.8	Length:	220
Score:	7.00 <td>Matches:</td> <td>7</td>	Matches:	7

```
CC ----- DUE TO A FRAMESHIFT. -----
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL; M58170; AAA62716.1; -.
DR EMBL; M33529; AAA30349.1; -.
DR PIR; JH0215; JH0215.
DR PIR; A38378; A38378.
DR PIR; A38379; A38379.
DR PDB; 1A0R; 16-FEB-99.
DR InterPro; IPR001200; Phosducin.
DR InterPro; IPR000063; Thioder.
DR Pfam; PF02114; Phosducin; 1.
DR PRINTS; PR00677; PHOSDUCIN.
KW Vision; Sensory transduction; Phosphorylation; 3D-structure.
FT MOD_RES 73 73
FT CONFLICT 44 44 H->P (IN REF. 3).
FT CONFLICT 238 239 TN->SK (IN REF. 3).
SQ SEQUENCE 245 AA; 28231 MW; 5CA621610401D550 CRC64;

Alignment Scores:
Pred. No.: 87.1 Length: 245
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x PHOS_BOVIN (1-245)

OY 679 TTACTGACCAATTCATATCA 699
DB 197 LeuLeuSerAsnPhelIeser 203
|||||
RESULT 16
PHOS_CANFA STANDARD: PRT; 245 AA.
AC 077560;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosducin (PHD) (33 kDa phototransducing protein).
GN PDC.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RX MEDLINE=98382516; PubMed=9714819;
RA Zhang Q., Acland G.M., Parshall C.J., Haskell J., Ray K.,
RA Aguirre G.D.;
RT "Characterization of canine photoreceptor phosducin cDNA and
RT identification of a sequence variant in dogs with photoreceptor
RT dysplasia.";
RL Gene 215:231-239(1998).
CC -1- FUNCTION: MAY PARTICIPATE IN THE REGULATION OF VISUAL
CC PHOTOTRANSDUCTION OR IN THE INTEGRATION OF PHOTORECEPTOR
CC METABOLISM.
CC -1- SUBUNIT: Forms a complex with the beta and gamma subunits of
CC the GTP-binding protein, transducin.
CC -1- SUBCELLULAR LOCATION: OUTER AND INNER SEGMENTS OF THE ROD CELLS
CC (BY SIMILARITY).
CC -1- DISEASE: DEFECTS IN PDC ARE THE CAUSE OF PHOTORECEPTOR DYSPLASIA
CC (PD); AN AUTOSOMAL RECESSIVE DISEASE OF MINIATURE SCHNAUZER DOGS
CC CAUSING RETINAL DEGENERATION.
CC -1- SIMILARITY: BELONGS TO THE PHOSDUCIN FAMILY.
```

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL; AF046874; AAC27249.1; -.
DR HSSP; P20942; 1B9X.
DR InterPro; IPR001200; Phosducin.
DR InterPro; IPR000063; Thioder.
DR Pfam; PF02114; Phosducin; 1.
DR PRINTS; PR00677; PHOSDUCIN.
KW Vision; Sensory transduction; Phosphorylation; Disease mutation.
FT MOD_RES 73 73
FT VARIANT 82 82 R->G (IN PD).
SQ SEQUENCE 245 AA; 28263 MW; CFA3779E2B7DD236 CRC64;

Alignment Scores:
Pred. No.: 87.1 Length: 245
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x PHOS_CANFA (1-245)

OY 679 TTACTGACCAATTCATATCA 699
DB 197 LeuLeuSerAsnPhelIeser 203
|||||
RESULT 17
CHER_BORBU STANDARD: PRT; 283 AA.
AC 051069;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chemotaxis protein methyltransferase (EC 2.1.1.80).
GN CHER OR BB0040.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RX MEDLINE=98065943; PubMed=9403685;
RC STRAIN=ATCC 35210 / B31;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uitterback T., Matthey L., McDonald L., Artlach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
CC -1- FUNCTION: METHYLATION OF THE MEMBRANE-BOUND METHYL-ACCEPTING
CC CHEMOTAXIS PROTEINS (MCP) TO FORM GAMMA-GLUTAMYL METHYL ESTER
CC RESIDUES IN MCP (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + protein L-glutamate
CC = S-adenosyl-L-homocysteine + protein L-glutamate methyl ester.
CC -1- SIMILARITY: CONTAINS 1 CHER-TYPE METHYLTRANSFERASE DOMAIN
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF001118; AAC66439.1; -
 DR HSSP: P07801; IBC5.
 DR TIGR: BB0040; -
 DR InterPro: IPR000780; Cher_Metranf.
 DR InterPro: IPR001601; Methyltransf.
 DR InterPro: IPR000051; SAM_bind.
 DR Pfam: PF01739; Cher; 1.
 DR Pfam: PF03705; Cher_N; 1.
 DR PRINTS: PR00996; CHERMTRFSE.
 DR SMART: SM00138; Metrc; 1.
 DR PROSITE: PS50123; CHER; 1.
 KM Transferrase; Methyltransferase; Chemotaxis; Complete proteome.
 FT DOMAIN 7 283 CHER-TYPE METHYLTRANSFERASE.
 SO SEQUENCE 283 AA; 33560 MW; 2F0E4E9C8BC3EE CRC64;
 Alignment Scores:
 Pred. No.: 84.8 Length: 283
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.86% Indels: 0
 DB: 1 Gaps: 0
 US-09-817-318-1 (1-780) x CHER_BORBU (1-283)
 QY 266 TTGACATTAACATCAACAAG 286
 Db 10 LeuasnleasnlethnLys 16
 RESULT 18
 PSTA_RHIL0 STANDARD; PRT; 283 AA.
 ID PSTA_RHIL0
 AC Q98FL4;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Phosphate transport system permease protein psta.
 GN PSTA OR ML13720
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=MAF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shlimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 CC -1- FUNCTION: PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
 CC FOR PHOSPHATE; PROBABLY RESPONSIBLE FOR THE TRANSLLOCATION OF THE
 CC SUBSTRATE ACROSS THE MEMBRANE (BY SIMILARTY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (By similarity).
 CC -1- SIMILARTY: BELONGS TO THE BINDING-PROTEIN-DEPENDENT TRANSPORT
 CC SYSTEM PERMEASE FAMILY. CYSTM SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: AP003002; BAB50553.1; -
 DR InterPro: IPR000515; BPD_transp.
 DR Pfam: PF00528; BPD_transp; 1.
 DR TIGRfams: TIGR00974; 3a0107s02c; 1.
 DR PROSITE: PS00402; BPD_TRANS_PNN_MEMBR; 1.
 KM Transport; Phosphate transport; Inner membrane;
 KW Complete proteome.
 FT TRANSMEM 20 42 POTENTIAL.
 FT TRANSMEM 70 92 POTENTIAL.
 FT TRANSMEM 105 127 POTENTIAL.
 FT TRANSMEM 137 156 POTENTIAL.
 FT TRANSMEM 188 210 POTENTIAL.
 FT TRANSMEM 254 276 POTENTIAL.
 SO SEQUENCE 283 AA; 30058 MW; 9CB19047AD82CC7F CRC64;
 Alignment Scores:
 Pred. No.: 84.8 Length: 283
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.86% Indels: 0
 DB: 1 Gaps: 0
 US-09-817-318-1 (1-780) x PSTA_RHIL0 (1-283)
 QY 532 ACATTGACTGACTTCGATTG 552
 Db 263 ThireuthrvalleuAlaleu 269
 RESULT 19
 MRW_BUCAP STANDARD; PRT; 312 AA.
 ID MRW_BUCAP
 AC O85295;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE S-adenosyl-methyltransferase mrw (EC 2.1.1.-).
 GN MRW.
 OS Buchnera aphidicola (subsp. Schizaphis graminum).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_TaxID=98794;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=98353428; PubMed=9688822;
 RA Thao M.L., Baumann P.;
 RT "Sequence analysis of a DNA fragment from Buchnera aphidicola (Aphid
 RT endosymbiont) containing the genes dapo-hrra-llvH-llvH-ftsl-ftsl-
 RT mure.";
 RL Curr. Microbiol. 37:214-216(1998).
 CC -1- FUNCTION: Exhibits a S-adenosyl-dependent methyltransferase
 CC activity (By similarity).
 CC -1- SIMILARTY: BELONGS TO THE MRW FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF060492; AAC32335.1; -
 DR InterPro: IPR002903; Bac_Metnfrse.
 DR Pfam: PF01795; Methyltransf.5; 1.
 DR ProDom: PD004685; Bac_Metnfrse; 1.
 DR TIGRfams: TIGR00006; UPF0117; 1.
 KM Transferrase; Methyltransferase.
 SO SEQUENCE 312 AA; 36039 MW; F5A7E7F803B31AE7 CRC64;
 Alignment Scores:
 Pred. No.: 83.3 Length: 312
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 1 Gaps: 0
US-09-817-318-1 (1-780) x MRAW_BUCAP (1-312)
OY 304 AGAATAGTCAGAACTTCATG 324
Db 247 Argilevallylsasphenet 253
RESULT 20
ID REBN_SALTY STANDARD; PRT; 314 AA.
AC P26403;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE O antigen biosynthesis rhamnosyltransferase rfbN (EC 2.4.1.-).
GN RFBN OR STM2085.
OS Salmonella typhimurium.
OC Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=91260454; PubMed=1710759;
RA Jiang X.-M., Neal B., Santiago F., Lee S.-J., Romana L.K., Reeves P.R.;
RT "Structure and sequence of the rfb (O antigen) gene cluster of
RT Salmonella serovar typhimurium (strain LT2).";
RL Mol. Microbiol. 5:695-713(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Portnoiiik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -1- CATALYTIC ACTIVITY: CDP-abequose + D-mannosyl-rhamnosyl-D-
CC galactose-1-diphospholipid -> CDP + abequosyl-D-mannosyl-rhamnosyl-
CC -1- galactose-1-diphospholipid.
CC -1- PATHWAY: O ANTIGEN BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X56793; CAA40127.1; -
DR EMBL: AE008792; AAL20969.1; -
DR PIR: S15311; S15311.
DR StcGene: SG10455; rfbN.
DR InterPro: IPR001173; Glycos_transf_2.
DR Pfam: PF00535; Glycos_transf_2; 1.
KW Lipopolysaccharide biosynthesis; Transferase; Glycosyltransferase;
KW Complete proteome.
SQ SEQUENCE 314 AA; 35508 MW; B15BA176EF2EDB7F CRC64;

Alignment Scores:
Pred. No.: 83.2 Length: 314
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0

DB: 1 Gaps: 0
US-09-817-318-1 (1-780) x RFBN_SALTY (1-314)
OY 267 AAGTCATTTGTGAGAGTAAAG 247
Db 138 LysSerIleVallySerLys 144
RESULT 21
ID HA10_MOUSE STANDARD; PRT; 328 AA.
AC P14428;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE H-2 class I histocompatibility antigen, K-Q alpha chain (H-2K(Q))
DE (Fragment).
GN H2-K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85206119; PubMed=3838967;
RA Morita T., Delarbre C., Kress M., Kourilsky P., Gachelin G.;
RT "An H-2K gene of the tw32 mutant at the T/t complex is a close parent
RT of an H-2Kq gene.";
RL Immunogenetics 21:367-383(1985).
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M14827; AAA39656.1; -
DR HSSP: P01901; 2VAA.
DR MGP: MGI:95904; H2-K.
DR InterPro: IPR003006; Iq_MHC.
DR InterPro: IPR003597; Iq_c1.
DR InterPro: IPR001039; MHC_I.
DR Pfam: PF00047; Iq; 1.
DR Pfam: PF00129; MHC_I; 1.
DR PRODOM: PD000050; MHC_I; 1.
DR SMART: SM00407; Iqcl; 1.
DR PROSITE: PS00290; Iq_MHC; 1.
KW MHC I; Transmembrane; Glycoprotein.
FT NON_TER 1 1
FT DOMAIN <1 71 EXTRACELLULAR ALPHA-1.
FT DOMAIN 72 163 EXTRACELLULAR ALPHA-2.
FT DOMAIN 164 255 EXTRACELLULAR ALPHA-3.
FT DOMAIN 256 265 CONNECTING PEPTIDE.
FT TRANSMEM 266 289
FT DOMAIN 290 328 CYTOPLASMIC TAIL.
FT DISULFID 82 145 BY SIMILARITY.
FT DISULFID 184 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 157 157 ADF1BD811BC37B4 CRC64;
SQ SEQUENCE 328 AA; 36855 MW; ADF1BD811BC37B4 CRC64;

Alignment Scores:
Pred. No.: 82.5 Length: 328
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0

DB: 1 Gaps: 0
 US-09-817-318-1 (1-780) x HA1Q_MOUSE (1-328)
 OY 625 ACTTCGATTATCCGCGCC 605
 ||||||||||||||||
 Db 309 ThersaspenseuSerLeupro 315

RESULT 22

HRDD_STRVF STANDARD: PRT: 332 AA.

ID HRDD_STRVF
 AC 060012:
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE RNA polymerase principal sigma factor hrdd.
 GN HRDD.

OS Streptomyces viridifaciens.
 OC Bacteria; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=48665;
 [1]

RP SEQUENCE FROM N.A.

RC STRAIN-MG456-hf10:
 RA U. W. Parry R.J.;

RT "Nucleotide sequence of the hrdd gene from the valani mycin
 producer, Streptomyces viridifaciens MG456-hf10.";
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

-1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
 ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
 THEN IS RELEASED.

CC -1- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: U60418; AAB03580.1; -
 CC HSSP: P00579; 1SIG.
 CC InterPro: IPR000943; Sigma_70.
 CC Pfam: PF00140; Sigma70; 1.
 CC PRINTS: PR00046; SIGMA70CT.
 CC PROSITE: PS00715; SIGMA70.1; 1.
 CC PROSITE: PS00716; SIGMA70.2; 1.
 CC KW Transcription regulation; Sigma factor; DNA-directed RNA polymerase;
 CC DNA-binding.

FT DOMAIN 124 137 POLYMERASE CORE BINDING (POTENTIAL).
 FT DNA_BIND 294 313 H-T-H MOTIF (BY SIMILARITY).
 FT SEQUENCE 332 AA; 37099 MW; E214BC68C57809AE CRC64;
 SQ

Alignment Scores:

Pred. No.: 82.3 Length: 332
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.92% Indels: 0
 DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x HRDD_STRVF (1-332)

OY 459 CAGAGAACTGCTCAATTATA 439

|||||
 Db 274 GlnarqThraLaserIelle 280

RESULT 23

MDHC_HUMAN STANDARD: PRT: 333 AA.

ID MDHC_HUMAN
 AC P40925;
 DT 01-FEB-1995 (Rel. 31, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Malate dehydrogenase, cytoplasmic (EC 1.1.1.37).
 GN MDH1 OR MDHA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 [1]

RP SEQUENCE FROM N.A.

RC TISSUE-Heart;
 RA Tanaka T., Inazawa J., Nakamura Y.;

RT "Molecular cloning and mapping of a human cDNA for cytosolic malate
 dehydrogenase (MDH1)."
 RL Genomics 32:128-130(1996).
 [2]

RP SEQUENCE FROM N.A.
 RC TISSUE-Heart;
 RA Lo A.S.Y., Waye M.M.Y.;

RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Placenta;
 RA Strausberg R.;

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 [4]
 RP SEQUENCE OF 167-180.
 RC TISSUE-Heart;
 RX MEDLINE=95203287; PubMed=7895732;

RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
 RT "The human myocardial two-dimensional gel protein database: update
 1994.";
 RL Electrophoresis 15:1459-1465(1994).
 CC -1- CATALYTIC ACTIVITY: (S)-malate + NAD(+) = oxaloacetate + NADH.

CC -1- SUBUNIT: HOMODIMER.
 CC -1- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: D55654; BAA09513.1; -
 CC EMBL: U20352; AAC16436.1; -
 CC DR EMBL: BC001484; AA01484.1; -
 CC HSSP: P11708; 4MDH.
 CC HSC-2DPAGE: P40925; HUMAN.
 CC DR Genew: HGNC:6970; MDH1.
 CC MIM: 154200; -
 CC InterPro: IPR001252; Mdh.
 CC InterPro: IPR001236; Ldh.
 CC Pfam: PF00056; Ldh; 1.
 CC Pfam: PF02866; Ldh.C; 1.
 CC PRODOM: PD003052; Mdh; 1.
 CC PROSITE: PS00068; MDH; 1.
 CC KW Oxidoreductase; Tricarboxylic acid cycle; NAD.
 CC Oxidoreductase; BY SIMILARITY.

FT INIT_MET 0
 FT ACT_SITE 158 158 PROTON-RELAY.
 FT BINDING 161 161 SUBSTRATE CARBOXYL GROUP.
 FT ACT_SITE 186 186 PROTON-RELAY.
 FT SEQUENCE 333 AA; 36295 MW; 888F9A79F6E5E888 CRC64;
 SQ

Alignment Scores:

Pred. No.: 82.3 Length: 333
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.92% Indels: 0


```

DB: 1 Gaps: 0
US-09-817-318-1 (1-780) x MDHC_HUMAN (1-333)
QY 437 GCAAGGAACTCAGACAGAGAA 417
    |||
Db 316 AlalysglutThrGlu 322

RESULT 24
MDHC_MOUSE
ID MDHC_MOUSE STANDARD: PRT: 333 AA.
AC P14152;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Malate dehydrogenase, cytoplasmic (EC 1.1.1.37).
GN MDH1 OR MOR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=68033094; PubMed=3312200;
RA Joh T., Takeshima H., Tsuzuki T., Setoyama C., Shimada K.,
RA Tanase S., Kuramitsu S., Kagamiyama H., Morino Y.;
RT "Cloning and sequence analysis of cDNAs encoding mammalian cytosolic
RT malate dehydrogenase. Comparison of the amino acid sequences of
RT mammalian and bacterial malate dehydrogenase."
RL J. Biol. Chem. 262:15127-15131(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C3H/He; TISSUE=Liver;
MDLINE=89011964; PubMed=3172222;
RA Setoyama C., Joh T., Tsuzuki T., Shimada K.;
RT Structural organization of the mouse cytosolic malate dehydrogenase
RT gene: comparison with that of the mouse mitochondrial malate
RT dehydrogenase gene."
RL J. Mol. Biol. 202:355-364(1988).
CC -1- CATALYTIC ACTIVITY: (S)-malate + NAD(+) = oxaloacetate + NADH.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: M29462; AAA39510.1; -
DR EMBL: M36084; AAA37423.1; -
DR PIR: S02654; DEMSMC.
DR HSSP: P11708; 4MDH.
DR SWISS-2DPAGE: P14152; MOUSE.
DR MGD: MGI:97051; Mor2.
DR InterPro: IPR001253; Mdh.
DR InterPro: IPR001236; Ldh.
DR Pfam: PF00056; Ldh; 1.
DR Pfam: PF00866; Ldh_C; 1.
DR ProDom: PD003052; Mdh; 1.
DR PROSITE: PS00068; MDH; 1.
KM Oxidoreductase; Tricarboxylic acid cycle; NAD.
FT INIT_MET 0
FT ACT_SITE 158 158 PROTON-RELAY.
FT BINDING 161 161 SUBSTRATE CARBOXYL GROUP.
FT ACT_SITE 186 186 PROTON-RELAY.
SO SEQUENCE 333 AA; 36346 MW; 01D5233640ED022A CRC64;
Alignment Scores: 82.3 Length: 333
Pred. No.:

```

```

Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 1 Gaps: 0
US-09-817-318-1 (1-780) x MDHC_MOUSE (1-333)
QY 437 GCAAGGAACTCAGACAGAGAA 417
    |||
Db 316 AlalysglutThrGlu 322

RESULT 25
CX3L_HUMAN
ID CX3L_HUMAN STANDARD: PRT: 355 AA.
AC P49238;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE CX3C chemokine receptor 1 (C-X3-C CMK-1) (CX3CR1) (Fractalkine
DE receptor) (GPR13) (V28) (Beta chemokine receptor-like 1) (CMK-BRL-1)
DE (CMKBLR1).
GN CX3CR1 OR GPR13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96011651; PubMed=7590284;
RA Raport C.J., Schweickart V.L., Eddy R.L., Jr., Shows T.B., Gray P.W.;
RT "The orphan G-protein-coupled receptor-encoding gene V28 is closely
RT related to genes for chemokine receptors and is expressed in lymphoid
RT and neural tissues."
RL Gene 163:295-299(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95374679; PubMed=7646814;
RA Combadiere C., Ahuja S.K., Murphy P.M.;
RT "Cloning, chromosomal localization, and RNA expression of a human
RT beta chemokine receptor-like gene."
RL DNA Cell Biol. 14:673-680(1995).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=98050927; PubMed=9390561;
RA Imai T., Hieshima K., Haskell C., Baba M., Nagira M., Nishimura M.,
RA Kakizaki M., Takagi S., Nomiyama H., Schall T.J., Yoshie O.;
RT "Identification and molecular characterization of fractalkine receptor
RT CX3CR1, which mediates both leukocyte migration and adhesion."
RL Cell 91:521-530(1997).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=98395093; PubMed=9726990.
RA Combadiere C., Salzwedel K., Smith E.D., Tiffany H.L., Berger E.A.,
RA Murphy P.M.;
RT "Identification of CX3CR1, A chemotactic receptor for the human CX3C
RT chemokine fractalkine and a fusion coreceptor for HIV-1."
RL J. Biol. Chem. 273:23799-23804(1998).
RN [5]
RP VARIANTS ALA-57; ILE-122; ILE-249 AND MET-280.
RX MEDLINE=20196025; PubMed=10731151;
RA Faure S., Meyer L., Costagliola D., Vaneensbergh C., Genin E.,
RA Autran B., Delfraissy J.-F., Mcdermott D.H., Murphy P.M., Debre P.,
RA Theodorou I., Combadiere C.;
RT "Rapid progression to AIDS in HIV+ individuals with a structural
RT variant of the chemokine receptor CX3CR1."
RL Science 287:2274-2277(2000).
CC -1- FUNCTION: RECEPTOR FOR THE CX3C CHEMOKINE AND MEDIATES
CC BOTH ITS ADHESIVE AND MIGRATORY FUNCTIONS. ACTS AS CO-RECEPTOR
CC WITH CD4 FOR HIV-1 VIRUS ENVELOPE PROTEIN (IN VITRO).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN LYMPHOID AND NEURAL TISSUES.
CC -1- DISEASE: INCREASED SUSCEPTIBILITY TO HIV INFECTION AND RAPID

```

PROGRESSION TO AIDS ARE ASSOCIATED WITH THE I-249/N-280 HAPLOTYPE.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U20350; AAA91763.1; -
 CC EMBL: U28934; AAA87032.1; -
 CC Genew: HGNC:2558; CX3CR1.
 CC MIM: 601470; -
 CC InterPro: IPR000276; GPCR_Rhodopsn.
 CC Pfam: PF00001; 7tm.1; 1.
 CC PRINTS: PR00237; GPCRHOOPS.
 CC PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 CC PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
 CC G-protein coupled receptor; Transmembrane; Polymorphism.
 FT DOMAIN 1 31
 FT TRANSMEM 32 59
 FT TRANSMEM 60 69
 FT TRANSMEM 70 90
 FT DOMAIN 91 103
 FT TRANSMEM 104 125
 FT DOMAIN 126 142
 FT TRANSMEM 143 167
 FT DOMAIN 168 195
 FT TRANSMEM 196 215
 FT DOMAIN 216 231
 FT TRANSMEM 232 256
 FT DOMAIN 257 273
 FT TRANSMEM 274 297
 FT DOMAIN 298 355
 FT DISULFID 102 175
 FT VARIANT 57 57
 FT
 FT VARIANT 122 122
 FT
 FT VARIANT 249 249
 FT
 FT VARIANT 280 280
 FT
 FT SEQUENCE 355 AA; 40396 MW; C59DC5F4C4312F22 CRC64;
 Alignment Scores:
 Pred. No.: 81.3 Length: 355
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.86% Indels: 0
 DB: 1 Gaps: 0
 US-09-817-318-1 (1-780) x C3X1_HUMAN (1-355)
 QY 704 TTAATTCCTCGTGTATATA 724
 Db 232 LeuileLeuValValille 238
 RESULT 26
 HAID_MOUSE STRAND: PRT; 368 AA.
 AC P01902;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE H-2 class I histocompatibility antigen, K-D alpha chain precursor
 DE (H-2K(D)).
 GN H2-K.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN
 RP SEQUENCE FROM N.A.
 RA Kvist S., Roberts L., Dobberstein B.;
 RT "Mouse histocompatibility genes: structure and organisation of a Kd
 RT gene.";
 RL EMBO J. 2:245-254(1983).
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/c;
 RX MEDLINE=83143367; PubMed=6298749;
 RA Lalanne J.-L., Delardre C., Gachet G., Kourilsky P.;
 RT "A cDNA clone containing the entire coding sequence of a mouse H-2Kd
 RT histocompatibility antigen.";
 RL Nucleic Acids Res. 11:1567-1577(1983).
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-NOD/LT;
 RA Girgis K.R., Capra D.J., Stroyanowski I.;
 RP Submitted (Mar-1995) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/c;
 RX MEDLINE=99086740; PubMed=9869916;
 RA Wang M., Stepkowski S.M., Hebert J.S., Tian L., Yu J., Kahan B.D.;
 RT "Nucleotide sequences of three H-2K and three H-2D complementary DNA
 RT clones coding mouse class I MHC heavy chain proteins.";
 RL Ann. Transplant. 1:26-31(1996).
 RN
 RP SEQUENCE OF 22-119.
 RX MEDLINE=81232882; PubMed=7018573;
 RA Kimball E.S., Nathanson S.G., Colligan J.E.;
 RT "Amino acid sequence of residues 1-98 of the K-2kb murine major
 RT histocompatibility alloantigen: comparison with H-2kb and H-2db
 RT reveals extensive localized differences.";
 RL Biochemistry 20:3301-3308(1981).
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: J00402; AAA39652.1; -
 CC EMBL: L36065; AAA89205.1; -
 CC EMBL: U47329; AAB17607.1; -
 CC PIR: A02204; HLM5KD.
 CC HSSP: P01899; 1B29.
 CC MGD: MGI:95904; H2-K.
 CC InterPro: IPR003006; Ig_MHC.
 CC InterPro: IPR003597; Ig_c1.
 CC InterPro: IPR01039; MHC_I.
 CC Pfam: PF00047; Ig_1.
 CC Pfam: PF00129; MHC_I; 1.
 CC ProDom: PD000050; MHC_I; 1.
 CC SMART: SM00407; IGc1; 1.
 CC PROSITE: PS00290; Ig_MHC; 1.
 CC MHC I; Transmembrane; Glycoprotein; Signal.
 FT CHAIN 1 21
 FT SIGNAL 22 368
 FT
 FT DOMAIN 22 111 H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN,
 FT DOMAIN 112 203 K-D ALPHA CHAIN.
 FT DOMAIN 204 295 EXTRACELLULAR ALPHA-1,
 FT DOMAIN EXTRACELLULAR ALPHA-2,
 FT DOMAIN EXTRACELLULAR ALPHA-3.

```
FT DOMAIN 296 305 CONNECTING PEPTIDE.
FT TRANSMEM 306 328
FT DOMAIN 329 368 CYTOPLASMIC TAIL.
FT DISULFID 122 185 BY SIMILARITY.
FT DISULFID 224 280 BY SIMILARITY.
FT CARBOHYD 107 107 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 135 135 O -> H (IN REF. 2).
SQ SEQUENCE 368 AA: 41490 MW: 4C83897C8CF76E5 CRC64;

Alignment Scores:
Pred. No.: 80.8 Length: 368
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x HAI1_MOUSE (1-368)
QY 625 ACTTCGATTATCCCTGCC 605
DB 349 ThrSerAspLeuSerLeuPro 355

RESULT 27
HAI1_MOUSE
ID HAI1_MOUSE STANDARD: PRT: 368 AA.
AC P03991: P03990.
DT 23-OCT-1986 (rel. 02, Created)
DT 01-JAN-1990 (rel. 13, Last sequence update)
DT 01-FEB-1996 (rel. 33, Last annotation update)
DE H-2 class I histocompatibility antigen, K-W28 alpha chain precursor.
GN H2-K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP MEDLINE=85206119; PubMed=3838967;
RA Morita T., Delafre C., Kress M., Kourilsky P., Gachelin G.;
RT "An H-2k gene of the tw32 mutant at the T/c complex is a close parent
of an H-2Kq gene."
RL Immunogenetics 21:367-383(1985).
RN [2]
RP SEQUENCE OF 258-368 FROM N.A.
RC STRAIN=SMR;
RA MEDLINE=84068207; PubMed=6689056;
RA Kress M., Glaros D., Khoury G., Jay G.;
RT "Alternative RNA splicing in expression of the H-2k gene."
RL Nature 306:602-604(1983).
CC -I- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -I- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M14825; AA39657.1; .
DR EMBL: X00172; CAA24997.1; .
DR PIR: A02197; A02197.
DR PIR: A02196; A02196.
DR HSSP: P01901: 2VAA.
DR MGD: MGI:95904; H2-K.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig-cl.
```

```
DR InterPro: IPR001039; MHC_I.
DR Pfam: PF00047; Ig: 1.
DR Pfam: PF00129; MHC_I; 1.
DR ProDom: PD000050; MHC_I; 1.
DR SMART: SM00407; Igcl; 1.
DR PROSITE: PS00290; IG_MHC; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 368
FT FT
FT DOMAIN 22 111 H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT DOMAIN 112 203 K-W28 ALPHA CHAIN.
FT DOMAIN 204 295 EXTRACELLULAR ALPHA-1.
FT TRANSMEM 296 305 EXTRACELLULAR ALPHA-2.
FT TRANSMEM 306 328 EXTRACELLULAR ALPHA-3.
FT DOMAIN 330 368 CONNECTING PEPTIDE.
FT DISULFID 122 185 CYTOPLASMIC TAIL.
FT DISULFID 197 185 BY SIMILARITY.
FT CARBOHYD 107 107 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 368 AA: 41103 MW: C07A9C0C458BA58F CRC64;

Alignment Scores:
Pred. No.: 80.8 Length: 368
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x HAI1_MOUSE (1-368)
QY 625 ACTTCGATTATCCCTGCC 605
DB 349 ThrSerAspLeuSerLeuPro 355

RESULT 28
HAI1_MOUSE
ID HAI1_MOUSE STANDARD: PRT: 369 AA.
AC P01901:
DT 21-JUL-1986 (rel. 01, Created)
DT 21-JUL-1986 (rel. 01, Last sequence update)
DT 15-JUL-1999 (rel. 38, Last annotation update)
DE H-2 class I histocompatibility antigen, K-B alpha chain precursor
(H-2K(b)).
GN H2-K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Weiss E., Golden L., Zakut R., Mellor A., Fahrner K., Kvist S.,
RA Flavel R.A.;
RT "The DNA sequence of the H-2K(b) gene: evidence for gene conversion
as a mechanism for the generation of polymorphism in
histocompatibility antigens."
RL EMBO J. 2:453-462(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/10;
RA MEDLINE=99086740; PubMed=9869916;
RA Wang M., Stepkowski S.M., Hebert J.S., Tian L., Yu J., Kahan B.D.;
RT "Nucleotide sequences of three H-2K and three H-2D complementary DNA
clones coding mouse class I MHC heavy chain proteins."
RL Ann. Transplant. 1:26-31(1996).
RN [3]
RP SEQUENCE OF 87-369 FROM N.A.
RA MEDLINE=82247837; PubMed=6954478;
RA Reyes A.A., Schold M., Itakura K., Wallace R.B.;
RT "Isolation of a cDNA clone for the murine transplantation antigen
H-2kb."
RL Proc. Natl. Acad. Sci. U.S.A. 79:3270-3274(1982).
```

```

RN [4]
RP SEQUENCE OF 22-367.
RX MEDLINE=82068960; PubMed=7306483;
RA Nephara H., Coligan J.E., Nathenson S.G.;
RT "Amino acid sequence of the carboxyl-terminal hydrophilic region of
  the H-2Kb MHC alloantigen. Completion of the entire primary structure
  of the H-2Kb molecule."
RL Biochemistry 20:5940-5945(1981).
RN [5]
RP REVISIONS TO 284; 296 AND 297.
RX MEDLINE=82247837; PubMed=6954478;
RA Coligan J.;
RT Unpublished results, cited by:
  Reyes A.A., Schold M., Itakura K., Wallace R.B.;
  Proc. Natl. Acad. Sci. U.S.A. 79:3270-3274(1982).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 22-295.
RX MEDLINE=92367003; PubMed=1323877;
RA Fremont D.H., Matsumura M., Stura E.A., Peterson P.A., Wilson I.A.;
RT "Crystal structures of two viral peptides in complex with murine MHC
  class I H-2Kb."
RL Science 257:919-927(1992).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 22-295 IN COMPLEX WITH CD8A.
RX MEDLINE=99021475; PubMed=9806638;
RA Kern P.S., Teng M.K., Smolyar A., Liu J.H., Liu J., Hussey R.E.,
  Speerl R., Chang H.-C., Reinherz E.L., Wang J.-H.;
RT "Structural basis of CD8 coreceptor function revealed by
  crystallographic analysis of a murine CD8alpha alpha ectodomain
  fragment in complex with H-2Kb."
RL Immunity 9:519-530(1998).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 22-295.
RX MEDLINE=99146375; PubMed=10023770;
RA Speir J.A., Abdel-Motal U.M., Jondal M., Wilson I.A.;
RT "Crystal structure of an MHC class I presented glycopeptide that
  generates carbohydrate-specific CTL."
RL Immunity 10:51-61(1999).
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
  THE IMMUNE SYSTEM.
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
  MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00400; AAA39648.1; -
DR EMBL: U47328; AAB17606.1; -
DR EMBL: V00746; CAA24119.2; -
DR EMBL: V00747; CAA24119.2; JOINED.
DR PIR: A02202; HELMSK8.
DR PDB: 2VAA; 20-JUN-96.
DR PDB: 2VAB; 20-JUN-96.
DR PDB: 2CKB; 09-SEP-98.
DR PDB: 1BOH; 19-AUG-98.
DR PDB: 1KBC; 09-FEB-99.
DR MGD; MG1:95904; H2-K.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_C1.
DR InterPro: IPR001039; MHC_I.
DR Pfam: PF00047; Ig_1.
DR Pfam: PF00129; MHC_I_1.
DR ProDom: PD000050; MHC_I_1.
DR SMART: SM00407; IGc1; 1.
DR PROSITE: PS00290; IG_MHC; 1.
KM MHC I; Transmembrane; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 21
FT CHAIN 22 369 H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN,

```

```

FT FT
FT DOMAIN 22 111 K-B ALPHA CHAIN.
FT DOMAIN 112 203 EXTRACELLULAR ALPHA-1.
FT DOMAIN 204 295 EXTRACELLULAR ALPHA-2.
FT DOMAIN 296 305 EXTRACELLULAR ALPHA-3.
FT TRANSMEM 306 328 CONNECTING PEPTIDE.
FT DOMAIN 329 369 CYTOPLASMIC TAIL.
FT DISULFID 122 185
FT CARBOHYD 107 107
FT CARBOHYD 197 197
FT CONFLICT 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 289 289 E -> D (IN REF. 4).
FT CONFLICT 334 334 E -> Q (IN REF. 4).
FT CONFLICT 364 364 N -> A (IN REF. 4).
FT CONFLICT 364 364 D -> P (IN REF. 4).
SO SEQUENCE 369 AA; 41302 MW; 3D2F125318193443 CRC64;

Alignment Scores:
Pred. No.: 80.8 Length: 369
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x HA1B_MOUSE (1-369)
QY 625 ACTCTGATTTATCCGCGCC 605
Db 350 ThrsraspleuSerleupro 356

RESULT 29
ID HA1K_MOUSE STANDARD: PRT; 369 AA.
AC P04223;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE H-2 class I histocompatibility antigen, K-K alpha chain precursor
DE (H-2K(K)).
GN H2-K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85087950; PubMed=6096831;
RA Arnold B., Burger H.G., Archibald A.L., Kvist S.;
RT "Complete nucleotide sequence of the murine H-2Kk gene. Comparison of
  three H-2K locus alleles."
RL Nucleic Acids Res. 12:9473-9487(1984).
[2]
RP SEQUENCE FROM N.A.
RX STRAIN=C3H;
RX MEDLINE=88060499; PubMed=3680952;
RA Watts S., Vogel J.M., Harriman W.D., Itoh T., Stauss H.J.,
  Goodenow R.S.;
RT "DNA sequence analysis of the C3H H-2K and H-2Dk loci. Evolutionary
  relationships to H-2 genes from four other mouse strains."
RL J. Immunol. 139:3878-3885(1987).
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
  THE IMMUNE SYSTEM.
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
  MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----

```

CC -I- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHRO

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L.

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Diez S.M.,
 RA Dodson K., Dopp L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Flischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.H., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-J., Ibegam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Moberly C., Morris J., Mostreli A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of *Drosophila melanogaster*.";
 RA Science 287:2185-2195(2000).
 RL [3]
 RP REVISIONS.
 RC STRAIN-Berkeley;
 RA Mistra S., Crosby M.A., Mathews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochank S.E., Smith C.D.,
 RA Tupy J.L., Bergman C.M., Bertram B.P., Carlson J.W., Geiniker S.E.,
 RA Clump M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
 RA Harris N.L., Krommiller B., Marshall B., Millburn G.H., Richter J.,
 RA Russo S., Seattle S.M.J., Smith E., Shu S., Smutlak F.,
 RA Whitfield E.J., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Mungall C.J., Lewis S.E.;
 RA "Annotation of *Drosophila melanogaster* genome.";
 RA Submitted (MAY-2002) to the EMBL/Genbank/DBU databases.
 RL [4]
 RP IDENTIFICATION AND TISSUE SPECIFICITY.
 RC MEDLINE-99166868; PubMed-10069338;
 RA Clyne P.J., Marr C.G., Freeman M.R., Lessing D., Kim J., Carlson J.R.;
 RA "A novel family of divergent seven-transmembrane proteins: candidate
 RT odorant receptors in *Drosophila*.";
 RL Neuron 22:327-338(1999).
 CC -i- FUNCTION: PROBABLE ROLE IN THE ODORANT RESPONSE, BEING AN ODORANT
 CC RECEPTOR.
 CC -i- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -i- TISSUE SPECIFICITY: EXPRESSED IN A SUBSET OF 17 OLFACTORY RECEPTOR
 CC NEURONS IN THE MAXILLARY PALP.
 CC -i- SIMILARITY: BELONGS TO FAMILY DR-OR OF G-PROTEIN COUPLED
 CC RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE003830; AAF58834.2; -;
 DR Flybase: FBgn0026388; O446a.
 DR InterPro: IPR004117; 7tm_6.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF02949; 7tm_6; 1.
 KW Transmembrane; G-protein coupled receptor; Olfaction; Glycoprotein;
 KW Multigene family.
 FT DOMAIN 1 37 EXTRACELLULAR (POTENTIAL).
 FT

FT TRANSMEM 38 58 1 (POTENTIAL).
 FT DOMAIN 59 65 2 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 66 86 2 (POTENTIAL).
 FT DOMAIN 87 127 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 128 148 3 (POTENTIAL).
 FT DOMAIN 149 170 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 171 191 4 (POTENTIAL).
 FT DOMAIN 192 255 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 256 276 5 (POTENTIAL).
 FT DOMAIN 277 287 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 288 308 6 (POTENTIAL).
 FT DOMAIN 309 354 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 355 375 7 (POTENTIAL).
 FT DOMAIN 376 385 CYTOPLASMIC (POTENTIAL).
 SQ SEQUENCE 385 AA; 44463 MW; BB062862B7A59310 CRC64;
 Alignment Scores:
 Pred. No.: 80.1 Length: 385
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.86% Indels: 0
 DB: 1 Gaps: 0
 US-09-817-318-1 (1-780) x 046A_DROME (1-385)
 OY 415 TTTTCTCTGTCAGTTCCTTT 435
 Db 362 PheSerSerValSerSerPhe 368
 ||||||||||||||||
 RESULT 32
 Y360_MYCGE
 ID Y360_MYCGE STANDARD; PRT; 411 AA.
 AC 049426;
 DT 01-NOV-1997 (rel. 35, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)
 DE Hypothetical protein MG360.
 GN MG360.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxId=2097;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 33530 / G-37;
 RX MEDLINE-96026346; PubMed-7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bolt K.F., Hu P.-C., Lucier T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RA "The minimal gene complement of *Mycoplasma genitalium*.";
 RL Science 270:397-403(1995).
 CC -i- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-Y FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U39717; AAC71585.1; -;
 DR TIGR: MG360; -;
 DR InterPro: IPR001126; UMOC_1like.
 DR Pfam: PF00817; IMS; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 411 AA; 47466 MW; F3FA0CA1E9018BAD CRC64;
 Alignment Scores:
 Pred. No.: 79.2 Length: 411

```

Score:              7.00          Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match:         2.86%       Indels: 0
DB:                  1          Gaps: 0

US-09-817-318-1 (1-780) x Y360_MYCGE (1-411)

OY 622 AACGTACTGCAATACAAAG 642
DB 326 LysleuLeuSerIleTnLys 332

RESULT 33
AMSI_ERWAM STANDARD: PRT: 415 AA.
ID AMSI_ERWAM
AC Q46637;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Amylovoran biosynthesis protein amsl.
GN AMSJ.
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95319333; PubMed=7596293;
RA Buger P., Gelder K.;
RT "Molecular analysis of the amsl operon required for exopolysaccharide
  synthesis of Erwinia amylovora."
RL Mol. Microbiol. 15:917-933(1995).
CC -1- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF AMYLOVORAN WHICH
  FUNCTIONS AS A VIRULENCE FACTOR.
CC -1- PATHWAY: Exopolysaccharide biosynthesis.
CC -1- SIMILARITY: TO E.COLI COLANIC ACID BIOSYNTHESIS PROTEIN WCAG.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC DR EMBL: X77921; CAA54888.1; -
CC KW Exopolysaccharide synthesis.
CC SQ SEQUENCE 415 AA; 46397 MW; 6295D941C55BD2BD CRC64;

Alignment Scores:
Pred. No.: 79 length: 415
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x AMSI_ERWAM (1-415)

OY 77 GCTGTAGCTATAATTATGAG 97
DB 341 AlavalaIleAsnTyGlu 347

RESULT 34
POTE_HAEIN STANDARD: PRT: 435 AA.
ID POTE_HAEIN
AC P44768;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putrescine-ornithine antiporter (Putrescine transport protein).
GN POTE OR HI0590.
OS Haemophilus influenzae.

```

```

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Rd / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McEneaney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Ueberlack T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
  Rd."
RL Science 269:496-512(1995).
CC -1- FUNCTION: PROBABLE PUTRESCINE-ORNITHINE ANTIPORTER
  (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
  (Potential).
CC -1- SIMILARITY: BELONGS TO THE ARCD/CADB/GERAB/LYSI FAMILY OF
  PERMEASES.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC DR EMBL: U32740; AAC22247.1; -
CC DR TIGR: HI0590; -
CC DR InterPro: IPR002293; AA/rel.prmaseel.
CC DR InterPro: IPR004841; Permease.
CC DR InterPro: IPR004754; R/O_antiport.
CC DR Pfam: PF00324; aa-permeases; 1.
CC DR TIGRFAMs: TIGR00905; ZA0302; 1.
CC KW Transport; Antiport; Amino-acid transport; Transmembrane;
  KW Inner membrane; Complete proteome.
CC FT TRANSMEM 8 POTENTIAL.
CC FT TRANSMEM 39 59 POTENTIAL.
CC FT TRANSMEM 95 115 POTENTIAL.
CC FT TRANSMEM 117 137 POTENTIAL.
CC FT TRANSMEM 148 168 POTENTIAL.
CC FT TRANSMEM 185 205 POTENTIAL.
CC FT TRANSMEM 224 244 POTENTIAL.
CC FT TRANSMEM 275 295 POTENTIAL.
CC FT TRANSMEM 320 340 POTENTIAL.
CC FT TRANSMEM 354 374 POTENTIAL.
CC FT TRANSMEM 386 406 POTENTIAL.
CC FT TRANSMEM 409 429 POTENTIAL.
CC SQ SEQUENCE 435 AA; 46348 MW; C4D992217975B5A8 CRC64;

Alignment Scores:
Pred. No.: 78.3 length: 435
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x POTE_HAEIN (1-435)

OY 233 ACAGCTTAGACCTTACTC 253
DB 329 ThrAlaLeuGInTrnLeu 335

RESULT 35
G6PI_LACIA

```

```

ID  G6PI_LACLA      STANDARD:      PRT:      447 AA.
AC  P81181;
DT  15-JUL-1998 (Rel. 36, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Glucose-6-phosphate isomerase (EC 5.3.1.9) (GPI) (Phosphoglucose
GN  isomerase) (PGI) (Phosphohexose isomerase) (PHI).
DE  PGI OR P6IA OR L2168.
OS  Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC  Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX  NCBI_TaxID=1360;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=IL1403;
RX  MEDLINE=21235186; PubMed=11337471;
RA  Bolotin A., Mincker P., Manger S., Jallion O., Malarme K.,
RA  Weisenbach J., Enllich S.D., Sorokin A.;
RT  "The complete genome sequence of the lactic acid bacterium Lactococcus
RT  lactis ssp. lactis IL1403.";
RL  Genome Res. 11:731-753(2001).
RN  [2]
RP  SEQUENCE OF 1-19.
RX  MEDLINE=97312580; PubMed=9169021;
RA  Nomura M., Nakajima I., Matsuzaki M., Kimoto H., Suzuki I., Aso H.;
RT  "The N-terminal sequence of Lactococcus lactis phosphoglucose
RT  isomerase purified by affinity chromatography differs from the other
RT  species.";
RL  Arch. Biochem. Biophys. 341:315-320(1997).
CC  -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC  phosphate.
CC  -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC  -1- SIMILARITY: BELONGS TO THE GPI FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: AE006445; AAK06266.1; -.
DR  HSSP: P13376; 2PGI.
DR  InterPro: IPR001672; G6P_Isomerase.
DR  Pfam: PF00342; PGI; 1.
DR  PRINTS: PR00662; G6PISOMERASE.
DR  PROSITE: PS00765; P_GLUCOSE_ISOMERASE_1; 1.
DR  PROSITE: PS00174; P_GLUCOSE_ISOMERASE_2; 1.
KW  Isomerase; Gluconeogenesis; Glycolysis; Complete proteome.
FT  INIT_MET 0
FT  ACT_SITE 310 BY SIMILARITY.
FT  ACT_SITE 424 BY SIMILARITY.
FT  ACT_SITE 424 BY SIMILARITY.
SQ  SEQUENCE 447 AA; 49463 MW; 40DD4FC7D106FD29 CRC64;

Alignment Scores:
Pred. No.: 78 Length: 447
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
Gaps: 0
DB: 1

US-09-817-318-1 (1-780) x G6PI_LACLA (1-447)
OY 391 TTATCCACTGTTTGTATAC 411
DB 94 LeuSerAsnSerPheValasn 100

RESULT 36
YKX8_YEAST STANDARD: PRT: 453 AA.
AC P34252;

```

```

DT  01-FEB-1994 (Rel. 28, Created)
DT  01-FEB-1994 (Rel. 28, Last sequence update)
DT  01-JUN-1994 (Rel. 29, Last annotation update)
DE  Hypothetical 52.3 kDa protein in HAP4-AAT1 intergenic region.
GN  YKL108W OR YKL463.
OS  Saccharomyces cerevisiae (Baker's yeast).
OC  Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC  Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX  NCBI_TaxID=4932;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=S288c;
RX  MEDLINE=94152173; PubMed=8109175;
RA  Cheret G., Pallier C., Valens M., Daiguan-Fornier B., Fukuhara H.,
RA  Bolotin-Fukuhara M., Sor F.;
RT  "The DNA sequence analysis of the HAP4-LAP4 region on chromosome XI
RT  of Saccharomyces cerevisiae suggests the presence of a second
RT  aspartate aminotransferase gene in yeast.";
RL  Yeast 9:1259-1265(1993).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: X71133; CAAS0449.1; -.
DR  EMBL: Z28108; CAAB1948.1; -.
DR  PIR: S37935; S37935.
DR  PIR: S39096; S39096.
DR  SCD: S0001591; YKL108W.
KW  Hypothetical protein.
SQ  SEQUENCE 453 AA; 52271 MW; 52B2969C97AAEAC8 CRC64;

Alignment Scores:
Pred. No.: 77.8 Length: 453
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
Gaps: 0
DB: 1

US-09-817-318-1 (1-780) x YKX8_YEAST (1-453)
OY 679 ATGAATTTATACCTATAAA 659
DB 97 MetAsnLeuSerProIleIys 103

RESULT 37
NU4M_HUMAN
ID  NU4M_HUMAN      STANDARD:      PRT:      459 AA.
AC  P03905;
DT  21-JUL-1986 (Rel. 01, Created)
DT  21-JUL-1986 (Rel. 01, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
GN  MTND4 OR ND4.
OS  Homo sapiens (Human).
OC  Mitochondrion.
OC  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  MEDLINE=81173052; PubMed=7219534;
RA  Anderson S., Bankier A.T., Barrell B.G., de Bruijn M.H.L.,
RA  Coulson A.R., Drouin J., Eperon I.C., Nierlich D.P., Roe B.A.,
RA  Sanger F., Schreier P.H., Smith A.J.H., Staden R., Young I.G.;
RT  "Sequence and organization of the human mitochondrial genome.";
RL  Nature 290:457-465(1981).
RN  [2]

```


RX MEDLINE-92315217; PubMed-1377597;
 RA Lu X., Walker T., Macmanus J.P., Seligy V.L.;
 RT "Differentiation of H7-29 human colonic adenocarcinoma cells
 RT correlates with increased expression of mitochondrial RNA: effects of
 RT trehalose on cell growth and maturation.";
 RL Cancer Res. 52:3718-3725(1992).
 RN [3]
 RP SEQUENCE OF 308-459 FROM N.A.
 RX MEDLINE-82242101; PubMed-6284948;
 RA Brown W.M., Prager E.M., Wang A., Wilson A.C.;
 RT "Mitochondrial DNA sequences of primates: tempo and mode of
 RT evolution.";
 RL J. Mol. Evol. 18:225-239(1982).
 RN [4]
 RP IDENTIFICATION OF PROTEIN.
 RX MEDLINE-85188293; PubMed-3921850;
 RA Chomay A., Mariotti P., Cleeter M.W.J., Ragan C.I., Matsuo-Yagi A.,
 RA Hatefi Y., Doolittle R.F., Atsardi G.;
 RT "Six unidentified reading frames of human mitochondrial DNA encode
 RT components of the respiratory-chain NADH dehydrogenase.";
 RL Nature 314:592-597(1985).
 RN [5]
 RP VARIANT LHON HIS-340.
 RX MEDLINE-89072713; PubMed-3201231;
 RA Wallace D.C., Singh G., Lott M.T., Hodge J.A., Schurr T.G.,
 RA Lezza A.M., Elsas L.J. II, Nikoskelainen E.K.;
 RT "Mitochondrial DNA mutation associated with Leber's hereditary optic
 RT neuropathy.";
 RL Science 242:1427-1430(1988).
 RN [6]
 RP CHARACTERIZATION OF VARIANT LHON HIS-340.
 RX MEDLINE-92070510; PubMed-1959619;
 RA Majander A., Huoponen K., Savontaus M.-L., Nikoskelainen E.,
 RA Wikstrom M.;
 RT "Electron transfer properties of NADH:ubiquinone reductase in the
 RT ND1/3460 and the ND4/11778 mutations of the Leber hereditary optic
 RT neuroretinopathy (LHON).";
 RL FEBS Lett. 292:289-292(1991).
 RN [7]
 RP VARIANT LHON HIS-340.
 RX MEDLINE-92070953; PubMed-1959931;
 RA Kornman B.A., Schuster H., Berringer T.A., Leo-Kottler B.;
 RT "Detection of the G to A mitochondrial DNA mutation at position 11778
 RT in German families with Leber's hereditary optic neuropathy.";
 RL Hum. Genet. 88:98-100(1991).
 RN [8]
 RP VARIANTS PRO-79; PRO-109; THR-132 AND THR-294.
 RX MEDLINE-92098084; PubMed-1757091;
 RA Marzuki S., Noer A.S., Lertit P., Thyagarajan D., Kapsa R.,
 RA Uthmanpol P., Byrne E.;
 RT "Normal variants of human mitochondrial DNA and translation products:
 RT the building of a reference data base.";
 RL Hum. Genet. 88:139-145(1991).
 RN [9]
 RP VARIANT MELAS ALA-109.
 RX MEDLINE-92359093; PubMed-1323207;
 RA Lertit P., Noer A.S., Jean-Francois M.J.B., Kapsa R., Dennett X.,
 RA Thyagarajan D., Lethlean K., Byrne E., Marzuki S.;
 RT "A new disease-related mutation for mitochondrial encephalopathy
 RT lactic acidosis and stroke-like episodes (MELAS) syndrome affects the
 RT ND4 subunit of the respiratory complex I.";
 RL Am. J. Hum. Genet. 51:445-448(1992).
 RN [10]
 RP VARIANT LDYT ILE-313.
 RX MEDLINE-96220221; PubMed-8644732;
 RA De Vries D.D., Went L.N., Bruyn G.W., Scholte H.R., Hofstra R.M.W.,
 RA Bolhuis P.A., Van Oost B.A.;
 RT "Genetic and biochemical impairment of mitochondrial complex I
 RT activity in a family with Leber hereditary optic neuropathy and
 RT hereditary spastic dystonia.";
 RL Am. J. Hum. Genet. 58:703-711(1996).
 RN [11]

US-09-817-318-1 (1-780) x NU4M_HUMAN (1-459)
OY 649 AACACTCTTGTATTGACA 629
|||||
DB 390 AsnIleThrLeuLeuThr 396

RESULT 38
NU4M_BRARE
ID NU4M_BRARE STANDARD; PRT; 460 AA.
AC P31394;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
GN NMND4 OR ND4.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxId=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB;
RX MEDLINE=21549095; PubMed=11691861;
RA Broughton R.E., Milam J.E., Roe B.A.;
RT "The complete sequence of the zebrafish (Danio rerio) mitochondrial
genome and evolutionary patterns in vertebrate mitochondrial DNA.";
RL Genome Res. 11:1958-1967(2001).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: AC024175; AAF74306.1; -
CC DR ZFIN: ZDB-GENE-011205-10; mtnd4.
CC DR InterPro: IPR001750; Oxidored_q1.
CC DR InterPro: IPR000280; Oxidored_q5_N.
CC DR Pfam: PF00361; oxidored_q1; 1.
CC DR Pfam: PF01059; oxidored_q5_N; 1.
CC KM Oxidoreductase; NAD; ubiquinone; Mitochondrion.
CC SQ SEQUENCE 460 AA; 51529 MW; 1615202B56513143 CRC64;

Alignment Scores:
Pred. No.: 77.5 Length: 460
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x NU4M_BRARE (1-460)
OY 273 TTACATCACAAGAGGTCC 293
|||||
DB 413 LeuThrSerGlnArgLysSer 419

RESULT 39
PRTG_RAT
ID PRTG_RAT STANDARD; PRT; 461 AA.
AC P31394;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vitamin-K dependent protein C precursor (EC 3.4.21.69)
DE (Autoproteolytic IIA) (Anticoagulant protein C) (Blood coagulation
DE factor XIV).
GN PROC.

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mistar; TISSUE=Liver;
RX MEDLINE=92329550; PubMed=1627650;
RA Okafuji T., Maekawa K., Nawa K., Marumoto Y.;
RT "The cDNA cloning and mRNA expression of rat protein C.";
RL Biochim. Biophys. Acta 1131:329-332(1992).
CC -1- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
CC REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VILLA
CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
CC -1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC and Villa.
CC -1- SUBUNIT: SYNTHESIZED AS A SINGLE CHAIN PRECURSOR, WHICH IS CLEAVED
CC INTO A LIGHT CHAIN AND A HEAVY CHAIN HELD TOGETHER BY A DISULFIDE
CC BOND. THE ENZYME IS THEN ACTIVATED BY THROMBIN, WHICH CLEAVES A
CC TETRADECAPEPTIDE FROM THE AMINO END OF THE HEAVY CHAIN; THIS
CC REACTION, WHICH OCCURS AT THE SURFACE OF ENDOTHELIAL CELLS, IS
CC STRONGLY PROMOTED BY THROMBOMODULIN.
CC -1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
CC -1- PTM: THE VITAMIN K-DEPENDENT, ENZYMATIC CARBOXYLATION OF SOME
CC GLU RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND CALCIUM.
CC -1- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO
CC ANOTHER SITE, BEYOND THE GLA DOMAIN. THIS GLA-INDEPENDENT BINDING
CC SITE IS NECESSARY FOR THE RECOGNITION OF THE
CC THROMBIN-THROMBOMODULIN COMPLEX.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: X64336; CAA45617.1; -
CC DR PIR: S18994; S18994.
CC DR PIR: S24312; S24312.
CC DR HSP: P04070; 1PCU.
CC DR MEROPS: S01.218; -
CC DR InterPro: IPR000152; Asx_hydroxyl.
CC DR InterPro: IPR001314; Chymotrypsin.
CC DR InterPro: IPR000561; EGF-like.
CC DR InterPro: IPR001881; EGF_CA.
CC DR InterPro: IPR002383; GLA_blood.
CC DR InterPro: IPR001284; Ser_protease_Try.
CC DR InterPro: IPR000294; VitK_dep_GLA.
CC DR Pfam: PF00008; EGF; 2.
CC DR Pfam: PF00089; trypsin; 1.
CC DR Pfam: PF00594; gla; 1.
CC DR PRINTS: PR00722; CHYMOTRYPSIN.
CC DR PRINTS: PR00001; GLABLOOD.
CC DR SMART: SM00179; EGF_CA; 1.
CC DR SMART: SM00001; EGF_like; 1.
CC DR SMART: SM00069; GLA; 1.
CC DR SMART: SM00020; TRYPSIN_SPEC; 1.
CC DR PROSITE: PS00010; ASX_HYDROXYL; 1.
CC DR PROSITE: PS00022; EGF_1; 1.
CC DR PROSITE: PS01186; EGF_2; 2.
CC DR PROSITE: PS01187; EGF_CA; 1.
CC DR PROSITE: PS00011; GLU_CARBOXYLATION; 1.
CC DR PROSITE: PS50240; TRYPSIN_DOM; 1.
CC DR PROSITE: PS00134; TRYPSIN_HIS; 1.
CC DR PROSITE: PS00135; TRYPSIN_SER; 1.
CC KW Blood coagulation; Glycoprotein; Serine protease;
KW Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT SIGNAL 1 32 BY SIMILARITY.

```
FT PROPEP 33 41 BY SIMILARITY.
FT CHAIN 42 196 PROTEIN C LIGHT CHAIN (BY SIMILARITY).
FT CHAIN 199 461 PROTEIN C HEAVY CHAIN (BY SIMILARITY).
FT PEPTIDE 199 212 ACTIVATION PEPTIDE (BY SIMILARITY).
FT SITE 212 212 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT DOMAIN 96 131 EGF-LIKE 1.
FT DOMAIN 135 175 EGF-LIKE 2.
FT DOMAIN 213 461 SERINE PROTEASE.
FT MOD_RES 47 47 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 48 48 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 55 55 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 57 57 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 60 60 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 61 61 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 66 66 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 67 67 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 70 70 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 112 112 HYDROXYLATION (BY SIMILARITY).
FT ACT_SITE 254 254 CHARGE RELAY SYSTEM.
FT ACT_SITE 300 300 CHARGE RELAY SYSTEM.
FT ACT_SITE 402 402 CHARGE RELAY SYSTEM.
FT DISULFID 58 63 BY SIMILARITY.
FT DISULFID 91 110 BY SIMILARITY.
FT DISULFID 100 105 BY SIMILARITY.
FT DISULFID 104 119 BY SIMILARITY.
FT DISULFID 121 130 BY SIMILARITY.
FT DISULFID 139 150 BY SIMILARITY.
FT DISULFID 146 159 BY SIMILARITY.
FT DISULFID 161 174 BY SIMILARITY.
FT DISULFID 182 320 INTERCHAIN (BY SIMILARITY).
FT DISULFID 239 255 BY SIMILARITY.
FT DISULFID 373 387 BY SIMILARITY.
FT DISULFID 398 426 BY SIMILARITY.
FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 461 AA: 51912 MW: 844CF93664EDACD5 CRC64:

Alignment Scores:
Pred. No.: 77.5 Length: 461
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x PRTC_RAT (1-461)
QY 454 TCCTGAATCACCACAACTG 474
DB 455 SerLeuysSerProLysLeu 461

RESULT 40
FMNL_HUMAN
ID FMNL_HUMAN STANDARD: PRT: 463 AA.
AC 095466: 096H10:
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Formin-like protein (Protein C17orf1).
DE Formin-like protein (Protein C17orf1).
GN FMNL OR C17ORF1B OR C17ORF1.
OS Homo sapiens (Human).
OC Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
OC Mammalia: Eutheria: Primates: Catarrhini: Homiini: Homo.
```

```
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99013435; PubMed=9799091.
RA Aronsson F.C., Magnusson P., Andersson B., Karsten S.L., Shibasaki Y.,
RA London C.L., Goate A.M., Brookes A.J.;
RT "The NIK protein kinase and C17orf1 genes: chromosomal mapping, gene
RT structures and mutational screening in frontotemporal dementia and
RT parkinsonism linked to chromosome 17."
RL Hum. genet. 103:340-345(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph. and Placenta;
RA Strausberg R.;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -I- TISSUE SPECIFICITY: Expressed in heart, brain, placenta, lung,
CC liver, skeletal muscle, kidney, and pancreas.
CC -I- SIMILARITY: CONTAINS 1 FORMIN HOMOLOG 2 (FH2) DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ008112; CA07870.1; -.
CC EMBL: BC001710; AAH01710.1; -.
CC EMBL: BC009000; AAH09000.1; -.
CC GeneW: HGNC:1212; FMNL.
CC MIM: 604656; -.
CC InterPro: IPR003104; FH2.
CC Pfam: Pf02181; FH2; 1.
CC SMART: SM00498; FH2; 1.
FT DOMAIN 1 426
FT CONFLICT 229 232 FH2.
SQ SEQUENCE 463 AA: 52433 MW: 17E9CA5A28BEA33E CRC64:

Alignment Scores:
Pred. No.: 77.5 Length: 463
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x FMNL_HUMAN (1-463)
QY 628 GTACTTCTGATTTATCCCTG 608
DB 457 ValThrSerAspLeuSerLeu 463

RESULT 41
AP54_YEAST
ID AP54_YEAST STANDARD: PRT: 475 AA.
AC 000776;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Clathrin coat assembly protein AP54 (Clathrin coat associated protein
DE AP54) (Golgi adaptor Ap-1 54 kDa protein) (HA1 54 kDa subunit)
DE (Clathrin assembly protein complex 1 medium chain).
GN APM1 OR YAP54 OR YPL259C OR P0394.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota: Fungi: Ascomycota: Saccharomycotina: Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92104180; PubMed=1761056;
RA Nakayama Y., Goebel M., O'Brien G.B., Lemmon S., Pingchang C.E.,
```

RA Kirchhausen T.;
 RT "The medium chains of the mammalian clathrin-associated proteins have
 RL a homolog in yeast."
 RN Eur. J. Biochem. 202:569-574(1991).
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288c / AB972;
 RX MEDLINE=97313271; PubMed=9169875;
 RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoerge W.,
 RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Barnes V.,
 RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
 RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
 RA Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,
 RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
 RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
 RA Hunkeler-Smith S., Hyman R., Johnston M., Kalman S., Klein K.,
 RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
 RA Marathe R., Messenguy F., Mewes H.-W., Mitterpatl S., Moestl D.,
 RA Mueller-Auer S., Namach A., Neutwich U., Oefner P., Pearson D.,
 RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
 RA Scherens B., Schramm S., Schroeder M., Sidcu A.M., Tettelin H.,
 RA Ustretarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
 RA Walsh S.V., Wambut R., Wang Y., Wedler E., Winnett E.,
 RA Zhong W.W., Zollner A., Vo D.H., Hani J.,
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI."
 RL Nature 387:103-105(1997).
 CC -1- FUNCTION: COMPONENT OF THE ADAPTOR COMPLEXES WHICH LINK CLATHRIN
 CC TO RECEPTORS IN COATED VESICLES. CLATHRIN-ASSOCIATED PROTEIN
 CC COMPLEXES ARE BELIEVED TO INTERACT WITH THE CYTOPLASMIC TAILS OF
 CC MEMBRANE PROTEINS, LEADING TO THEIR SELECTION AND CONCENTRATION.
 CC AP57 IS PROBABLY A SUBUNIT OF THE GOLDI MEMBRANE ADAPTOR.
 CC -1- SUBUNIT: ASSEMBLY SUBUNIT OF THE GOLDI MEMBRANE ADAPTOR
 CC COMPOSED OF TWO LARGE CHAINS (GAMMA AND BETA'), A MEDIUM CHAIN
 CC (AP54) AND A SMALL CHAIN (AP19).
 CC -1- SUBCELLULAR LOCATION: COATED VESICLE.
 CC -1- SIMILARITY: BELONGS TO THE ADAPTOR COMPLEXES MEDIUM SUBUNITS
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X60288; CAA4282.1; -
 DR EMBL: Z73615; CAA97989.1; -
 DR PIR: S17028; S17028.
 DR TRANSFAC: T03713; -
 DR SGD: S0006180; APM1.
 DR InterPro: IPR001392; Clathrm_med.
 DR Pfam: PF00928; Adap_comp_sub.1.
 DR PRINTS: PR00314; CLATHRINADPT.
 DR PROSITE: PS00990; CLAT_ADAPTOR_M_1.1.
 DR PROSITE: PS00991; CLAT_ADAPTOR_M_2.1.
 KW Coated pits; Phosphorylation.
 CC
 FT CONFLICT 214 214 M -> I (IN REF. 1).
 FT CONFLICT 216 216 D -> H (IN REF. 1).
 FT CONFLICT 222 222 N -> K (IN REF. 1).
 FT CONFLICT 433 433 P -> R (IN REF. 1).
 FT CONFLICT 440 440 I -> M (IN REF. 1).
 FT CONFLICT 450 450 MISSING (IN REF. 1).
 FT CONFLICT 450 450 MISSING (IN REF. 1).
 SQ SEQUENCE 475 AA: 53873 MW: C6B11153B845921C CRC64;

Alignment Scores:

Pred. No.: 77.1 Length: 475
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.92% Indels: 0
 DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x AP54_YEAST (1-475)
 Oy 283 TGTGATGTATATGTTCAAGTC 263
 ||||||||||||||||||
 Db 318 CysAspValAsnValGlnVal 324
 RESULT 42
 YDE5_SCHPO
 ID YDE5_SCHPO STANDARD; PRT; 486 AA.
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable peptidase Cl2B10.05 (EC 3.4.-.-).
 GN SPAC12B10.05.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RP SEQUENCE FROM N.A.
 RC STRAIN-972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor A., Cronin A., Davis P., Felwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymptre B.,
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach E., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler R., Purnelle B.,
 RA Goffeau A., Gadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Kochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Pausen P.;
 RT "The genome sequence of Schizosaccharomyces pombe."
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M24B.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: Z70721; CAA4695.1; -
 DR HSSP: P15034; IAI6.
 DR MEROPS: M24_026; -
 DR InterPro: IPR000994; Peptidase_M24.
 DR InterPro: IPR001131; Xaa-Pro-peptids.
 DR Pfam: PF00557; Peptidase_M24.1.
 DR PROSITE: PS00491; PROLINE_PEPTIDASE; 1.
 KW Hypothetical protein; Hydrolase.
 SQ SEQUENCE 486 AA: 55140 MW: 113B6FD4758D989B CRC64;

Alignment Scores:

Pred. No.: 76.8 Length: 486
 Score: 7.00 Matches: 7

Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 2.92%
 DB: 1
 Gaps: 0

US-09-817-318-1 (1-780) x YDE5_SCHPO (1-486)

OY 725 ATATTACACGAGCAATTA 705
 ||||||||||||||||
 Db 52 ILEuclnProglYgluleu 58

RESULT 43

VGA_BPPHX STANDARD; PRT; 513 AA.
 AC P03631:
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE A and A* proteins (GPA).
 GN A.
 OS Bacteriophage phi-X174.
 OC Viruses; ssDNA viruses; Microviridae; Microvirus.
 OX NCBI_TaxID=10847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=77171175; PubMed=870828;
 RA Sanger F., Air G.M., Barrell B.G., Brown N.L., Coulson A.R.,
 RA Fiddes J.C., Hutchison C.A. III, Slocombe P.M., Smith M.;
 RT "Nucleotide sequence of bacteriophage phi X174 DNA.";
 RL Nature 265:687-695(1977).
 RN [2]
 RP REVISIONS.
 RX MEDLINE=79091185; PubMed=731693;
 RA Sanger F., Coulson A.R., Friedmann T., Air G.M., Barrell B.G.,
 RA Brown N.L., Fiddes J.C., Hutchison C.A. III, Slocombe P.M., Smith M.;
 RT "The nucleotide sequence of bacteriophage phiX174.";
 RL J. Mol. Biol. 125:225-246(1978).
 RN [3]
 RP FUNCTION OF A AND A*
 RX MEDLINE=80027245; PubMed=158588;
 RA Ikeda J., Ydelevich A., Shimamoto N., Hurwitz J.;
 RT "Role of polymeric forms of the bacteriophage phi X174 coded gene A
 RT protein in phi XRFI DNA cleavage.";
 RL J. Biol. Chem. 254:9416-9428(1979).
 RN [4]
 RP ACTIVE SITE TYROSINE.
 RX MEDLINE=86232554; PubMed=2940511;
 RA van Mansfield A.D.M., van Teeffelen H.A.A.M., Baas P.D., Jansz H.S.;
 RT "Two juxtaposed tyrosyl-OH groups participate in phi X174 gene A
 RT protein catalysed cleavage and ligation of DNA.";
 RL Nucleic Acids Res. 14:4229-4238(1986).
 CC -1- FUNCTION: THE A PROTEIN IS A SPECIFIC ENDONUCLEASE THAT CLEAVES
 CC THE VIRAL STRAND OF SUPERSTWISTED, CLOSED CIRCULAR DNA AT A UNIQUE
 CC SITE IN THE A GENE. THE A PROTEIN ALSO CAUSES RELAXATION OF
 CC SUPERSTWISTED DNA AND FORMS A COMPLEX WITH VIRAL DNA THAT HAS A
 CC DISCONTINUITY IN GENE A OF THE VIRAL STRAND
 CC -1- FUNCTION: GENE A* PROTEIN IS FORMED FROM A NATURAL INITIATION SITE
 CC WITHIN THE A GENE. IT BINDS TO DOUBLE-STRANDED DNA AND PREVENTS
 CC THEIR HYDROLYSIS BY NUCLEASES.
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM MEMBRANE (PROBABLY).
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: V01128; CAA24351.1; -
 CC EMBL: J02482; AAA32570.1; -
 CC EMBL: J02482; AAA32570.1; ALT_INIT.
 CC PIR: A04239; ZABPA.
 CC Hydrolase; Nuclease; Endonuclease; DNA-binding; Zinc-finger.

FT CHAIN 1 513 PROTEIN A (BY SIMILARITY WITH G4).
 FT CHAIN 173 513 PROTEIN A* (BY SIMILARITY WITH G4).
 FT ZN_FING 246 268 POTENTIAL.
 FT BINDING 347 347 TO DNA (WHEN NICKING OCCURS).
 SO SEQUENCE 513 AA; 58718 MW; EF2D4DC9E3BC8072 CRC64;

Alignment Scores:

Pred. No.: 76 Length: 513
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.92% Indels: 0
 DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x VGA_BPPHX (1-513)

OY 234 GTCACACATTTTGTAGAAA 214
 ||||||||||||||||
 Db 132 ValThrGlnPheCysArgGlys 138

RESULT 44

SC59_YEAST STANDARD; PRT; 519 AA.
 AC P20048:
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Dolichol kinase (EC 2.7.1.108).
 GN SEC59 OR YMR013C OR YMR270.17C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89261723; PubMed=2657387;
 RA Bernstein M., Kepes F., Schekman R.;
 RT "Sec59 encodes a membrane protein required for core glycosylation in
 RT Saccharomyces cerevisiae.";
 RL Mol. Cell. Biol. 9:1191-1199(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Deylin K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP CHARACTERIZATION.
 RX STRAIN=S288C;
 RC MEDLINE=92357761; PubMed=1323123;
 RA Heller L., Orlean P., Adair W.L. Jr.;
 RT "Saccharomyces cerevisiae sec59 cells are deficient in dolichol
 RT kinase activity.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:7013-7016(1992).
 CC -1- FUNCTION: INVOLVED IN THE SYNTHESIS OF N-LINKED AND O-LINKED
 CC WHICH IS REQUIRED IN THE SYNTHESIS OF N-LINKED AND O-LINKED
 CC OLIGOSACCHARIDES AND FOR THAT OF GPI ANCHORS. IT IS REQUIRED FOR
 CC SPORE GERMINATION. HAS AN ESSENTIAL ROLE IN CELLULAR METABOLISM.
 CC -1- CATALYTIC ACTIVITY: cnp + dolichol = CDP + dolichyl phosphate.
 CC -1- PATHWAY: Glycosylation.
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM MEMBRANE (PROBABLY).
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M25779; AAA35033.1; -
 CC EMBL: 248613; CAA88530.1; -
 CC PIR: J00124; J00124.
 CC SGD: S0004615; SEC59.

```

DR      InterPro: IPR002741; DUF56.
DR      Pfam: PF01879; DUF56; 1.
KW      Endoplasmic reticulum; Membrane; Transferase; Kinase.
FT      CARBOHYD 326 326 N-LINKED (GLCNAC...) (POTENTIAL).
FT      CARBOHYD 415 415 N-LINKED (GLCNAC...) (POTENTIAL).
SQ      SEQUENCE 519 AA; 38906 MW; 2C75B6F175BBFF06 CRC64;

Alignment Scores:
Pred. No.: 75.8 Length: 519
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x SC59_YEAST (1-519)
QY      667 GTGATAATTCATTACTGAGC 687
Db      185 ValIleasnserleuSer 191

RESULT 45
VGA_BPS13
ID      VGA_BPS13 STANDARD; PRT; 522 AA.
AC      P07928;
DT      01-AUG-1988 (Rel. 08, Created)
DT      01-AUG-1988 (Rel. 08, Last sequence update)
DE      15-JUL-1998 (Rel. 36, Last annotation update)
DE      A', A and A* proteins (GPA).
GN      A.
OS      Bacteriophage S13.
OC      Viruses; ssDNA viruses; Microviridae; Microvirus.
OX      NCBI_TaxID=10844;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=86165869; PubMed=3007293;
RA      Lau P.C.K., Spencer J.H.;
RT      "Nucleotide sequence and genome organization of bacteriophage S13
RT      DNA."
RL      Gene 40:273-284(1985).
RN      [2]
RP      SEQUENCE OF 269-333 FROM N.A.
RX      MEDLINE=78011668; PubMed=909772;
RA      Grosveld F.G., Spencer J.H.;
RT      "The nucleotide sequence of two restriction fragments located in the
RT      gene AB region of bacteriophage S13."
RL      Nucleic Acids Res. 4:2235-2252(1977).
CC      -!- FUNCTION: THE A PROTEIN IS A SPECIFIC ENDONUCLEASE THAT CLEAVES
CC      THE VIRAL STRAND OF SUPERTWISTED, CLOSED CIRCULAR DNA AT A UNIQUE
CC      SITE IN THE A GENE. THE A PROTEIN ALSO CAUSES RELAXATION OF
CC      SUPERTWISTED DNA AND FORMS A COMPLEX WITH VIRAL DNA THAT HAS A
CC      DISCONTINUITY IN GENE A OF THE VIRAL STRAND.
CC      -!- FUNCTION: GENE A* PROTEIN IS FORMED FROM A NATURAL INITIATION SITE
CC      WITHIN THE A GENE. IT BINDS TO DOUBLE-STRANDED DNA AND PREVENTS
CC      THEIR HYDROLYSIS BY NUCLEASES.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: M14428; AAA32581.1; -
DR      EMBL: M14428; AAA32582.1; -
DR      EMBL: M14428; AAA32583.1; -
DR      PIR: J50450; J50450.
KW      Hydrolyase; Nuclease; Endonuclease; DNA-binding; Zinc-finger.
FT      CHAIN 1 522 PROTEIN A'.
FT      CHAIN 10 522 PROTEIN A.
FT      CHAIN 182 522 PROTEIN A*.
FT      ZN_FING 255 277 POTENTIAL.

```

```

FT      BINDING 356 356 TO DNA (WHEN NICKING OCCURS)
FT      SQ      SEQUENCE 522 AA; 59882 MW; CE9DE9AC679403A2 CRC64;
SQ      SEQUENCE 522 AA; 59882 MW; CE9DE9AC679403A2 CRC64;

Alignment Scores:
Pred. No.: 75.8 Length: 522
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x VGA_BPS13 (1-522)
QY      234 GTCACACAAATTTGTAGAAA 214
Db      141 ValThrGlnphecysarglys 147

```

Search completed: February 4, 2003, 07:18:19
Job time : 26.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 4, 2003, 07:17:33 ; Search time 43 Seconds
(without alignments)
3487.667 Million cell updates/sec

```
Title: US-09-817-318-1
Perfect score: 245
Sequence: 1 aatracctgttctcttaag.....taagcgcaaaaaaaaaa 780
```

Scoring table:	
OLIGO	
Xgapop 60.0 ,	Xgapext 60.0
Ygapop 60.0 ,	Ygapext 60.0
Fgapop 6.0 ,	Fgapext 7.0
Delop 6.0 ,	Delext 7.0

Searched: 283224 seqs, 96134422 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565536

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 45 summaries

Command line parameters:

```

MODEL=firmat2p.model /usr-xlp
-o/cgna2.1/USPRO.spool/USDB17318/runat_04022003_071723_6638/app_query.fasta.1.965
-DB=PIR.73 -QEXT=fastcan -SUFFIX=rrp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=0110 -TRANS=human40.0d1 -LIST=45
-DUPLICATION=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=45 -MODE=LOCAL -OUTFWT=pto
-NORR=ext -HEAPSITE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=0509817318 GCN=1.148 -etunal_04022003_071723_6638 -NCPU=6 -ICPU=3
-NO_XLXPY -NO_MMAR -LARGOQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=10 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

```

Database :

```
PIR_73: *
1:  pir1: *
2:  pir2: *
3:  pir3: *
4:  pir4: *
```

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	8	3.3	145	2	G81263	probable periplasm
2	8	3.3	235	2	T04451	hypothetical protein
3	8	3.3	266	2	C75613	probable molybdenum
4	8	3.3	341	2	T46153	hypothetical protein
5	8	3.3	451	2	T01400	translation elonga
6	8	3.3	528	2	S67624	ASMA protein - yea
7	8	3.3	607	2	C64420	oligoendopeptidase
8	8	3.3	866	2	C97662	hypothetical protein
9	8	3.3	866	2	AF2886	conserved hypot
10	8	3.3	963	2	H85015	hypothetical prote
11	8	3.3	1613	2	E82193	conserved hypot
12	7	2.9	20	2	A60897	class I histocompa
13	7	2.9	31	2	S58771	NADH dehydrogenas
14	7	2.9	76	2	E64453	hypothetical protei

15	hypothetical prote	7	2.9	91	2	T43310
16	MHC class I histoc	7	2.9	102	2	A02196
17	thioredoxin 1 - SL	7	2.9	105	1	A45264
18	MHC class I histoc	7	2.9	111	2	A02197
19	neurotoxin homolog	7	2.9	134	2	I48639
20	hypothetical prote	7	2.9	146	2	T13311
21	hypothetical prote	7	2.9	147	2	T18586
22	hypothetical prote	7	2.9	153	2	G87632
23	MHC class I antigen	7	2.9	164	2	I58201
24	hypothetical prote	7	2.9	168	2	D90066
25	transcripton repr	7	2.9	172	2	A35145
26	DNA-directed RNA p	7	2.9	173	1	JT0302
27	acetylcholinestera	7	2.9	173	2	A39022
28	hypothetical prote	7	2.9	174	2	S45548
29	hypothetical prote	7	2.9	175	2	D75514
30	hypothetical prote	7	2.9	199	2	S14981
31	extensin class I (7	2.9	206	2	T31585
32	hypothetical prote	7	2.9	211	2	T17228
33	hypothetical prote	7	2.9	219	2	T45997
34	probable zinc fing	7	2.9	219	2	S59134
35	hypothetical prote	7	2.9	220	2	S40931
36	GTP cyclohydrolase	7	2.9	222	2	A10781
37	hypothetical prote	7	2.9	224	2	D88836
38	two-component resp	7	2.9	231	2	G70000
39	two-component resp	7	2.9	231	2	G84138
40	hemoglobin linker	7	2.9	235	2	S43513
41	two-component resp	7	2.9	237	2	D70032
42	probable RNA polym	7	2.9	238	2	A81422
43	phosducin, retinal	7	2.9	245	1	A38379
44	hypothetical prote	7	2.9	245	2	G81297
45	vrio protein - Dic	7	2.9	251	2	T17397

ALIGNMENTS

```

RESULT 1
GB1263
probable periplasmic protein Cj1666c [imported] - Campylobacter jejuni (strain NCTC 11168)
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: GB1263
R:Packhill, J.; Trenn, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Parkhill, J.; Mungall, K.M.; VanVleet, A.; Whitehead, S.; Barrow, P.;
C.W.: Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrow, P.;
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: GB1263
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-145 <PAR>
A:Cross-references: GB:AL139079; GB:AL111168; NID:96968971; PIDN:CAE73653.1; PID:96968971;
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj1666c
C:Superfamily: Campylobacter jejuni probable periplasmic protein Cj1666c

Alignment Scores:
Pred. NO.: 13.5 Length: 145
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x GB1263 (1-145)

QY 651 TATGGGATTTTATAGGTGATATAA 674
|||||
Db 136 TyrglyTyrPhelIeGlyAspIys 143

RESULT 2
T04451
hypothetical protein F4D11.60 - Arabidopsis thaliana

```

C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 26-Aug-1999
C:Accession: T04451
R:Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansgorge, W.; Hohnsels, J.; Mewes, H.W.
submitted to the Protein Sequence Database, April 1998
A:Reference number: Z15360
A:Accession: T04451
A:Molecule type: DNA
A:Residues: 1-235 <BEV>
A:Cross-references: EMBL:AL022537
A:Experimental source: cultivar Columbia; BAC clone F4D11
C:Genetics:
A:Map position: 4
A:Insertions: 66/2; 104/3; 196/3
A:Note: F4D11.60
C:Superfamily: Arabidopsis hypothetical protein F4D11.60

Alignment Scores:
Pred. No.: 12.9 Length: 235
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x T04451 (1-235)

QY 482 CATAAATCAGTTGGCTGATTTC 459
|||||
DB 21 HistylleSerleuGlyAspHe 28

RESULT 3
C75613
probable molybdenum ABC transporter, permease protein - Deinococcus radiodurans (strain
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: C75613
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: C75613
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-266 <WHI>
A:Cross-references: GB:AE001862; GB:AE001825; NID:g6460468; PIDN:AAF12201.1; PID:g646046
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0168
A:Map position: 2
C:Superfamily: maltose transport protein malG

Alignment Scores:
Pred. No.: 12.7 Length: 266
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x C75613 (1-266)

QY 646 ATAAGTTGTTATGACAGTAAC 623
|||||
DB 61 IllethleuLeuThrValThr 68

RESULT 4
T46153
hypothetical protein T3A5.140 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 18-Feb-2000

C:Accession: T46153; T08397
R:Boecker, H.; Mewes, H.W.; Mayer, K.F.X.; Lemcke, K.; Schueller, C.; Quetier, F.; S
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z23024
A:Accession: T46153
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-341 <BL0>
A:Cross-references: EMBL:AL132979
A:Experimental source: cultivar Columbia; BAC clone T3A5
R:Quetier, F.; Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Salan
submitted to the Protein Sequence Database, May 1999
A:Reference number: Z16409
A:Accession: T08397
A:Molecule type: DNA
A:Residues: 1-341 <QUE>
A:Cross-references: EMBL:AL049862; GSPDB:GN00061; ATSP:FI8B3.40
A:Experimental source: cultivar Columbia; BAC clone FI8B3
C:Genetics:
A:Gene: ATSP:FI8B3.40
A:Map position: 3
A:Note: T3A5.140

Alignment Scores:
Pred. No.: 12.4 Length: 341
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x T46153 (1-341)

QY 98 AGTAACTTATTTATATATATCC 121
|||||
DB 3 SerlysheilleuThrleuSer 10

RESULT 5
T01400
translation elongation factor EF-Tu precursor, mitochondrial - Arabidopsis thaliana
N:Alternate names: protein T419.19
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 02-Feb-2001
C:Accession: T01400; S62701
R:Parnell, L.D.; Gnoj, L.; de la Bastide, M.; Hameed, A.; Habermann, K.; Schutz, K.;
submitted to the EMBL data library, May 1998
A:Description: Genomic sequence of BAC T419 from Arabidopsis thaliana, chromosome IV,
A:Reference number: Z14314
A:Accession: T01400
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-454 <PAR>
A:Cross-references: EMBL:AF069442; NID:g3242970; PIDN:MAC79113.1; PID:g3924612
A:Experimental source: cultivar Columbia
R:Kuhlman, P.; Palmer, J.D.
Plant Mol. Biol. 29, 1057-1070, 1995
A:Title: Isolation, expression, and evolution of the gene encoding mitochondrial elon
A:Reference number: S62701; MUID:96145515; PMID:855448
A:Accession: S62701
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 'MEEPVREDIDLRVSSDTIGW', 8, 'EFAAA', 11-309, 'CST', 313-454 <RUH>
A:Cross-references: EMBL:X89227; NID:g1149570; PIDN:CAA61511.1; PID:g1149571
C:Genetics:
A:Gene: tu1m
A:Map position: 4
A:Genome: nuclear
A:Insertions: 64/2; 89/3; 120/3; 145/3; 180/3; 211/1; 279/3; 307/3; 330/3; 351/3; 432/1
A:Note: T419.19
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom
C:Keywords: GTP binding; mitochondrion; nucleotide binding; P-loop; protein biosynthe
F1-55/Domain: transit peptide (mitochondrion) #status predicted <TRN>
F1-56-454/Product: translation elongation factor EF-Tu #status predicted <MAT>

RESULT 9
AF2886
conserved hypothetical protein Atu2521 [imported] - *Agrobacterium tumefaciens* (strain C58)
C:Species: *Agrobacterium tumefaciens*
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AF2886
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Genthner, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan,
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AF2886
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-866 <KUR>
A:Cross-References: GB:AE008688; PIDN:AAL43508.1; PID:g17741016; GSPDB:GN00186
C:Genetics:
A:Experimental source: strain C58 (Dupont)
A:Gene: Atu2521
A:Map position: circular chromosome

Alignment Scores:
Pred. No.: 11.3 Length: 866
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.27% Indels: 0
DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x AF2886 (1-866)

OY 428 GTTCCTTGGCTATTAATGACGA 451
|||||
Db 376 ValProLeuProlleileGuaLa 383

RESULT 10
H85015
hypothetical protein AT4g01210 [imported] - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C:Accession: H85015
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*.
A:Reference number: AB5001; MUID:20083488; PMID:10617198
A:Accession: H85015
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-963 <STO>
A:Cross-References: GB:NC_001268; NID:g7267618; PIDN:CA80930.1; GSPDB:GN00140
A:Gene: AT4g01210
A:Map position: 4
C:Superfamily: *Arabidopsis thaliana* hypothetical protein A_IG002N01.24

Alignment Scores:
Pred. No.: 11.2 Length: 963
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x H85015 (1-963)

OY 348 AGAAAGACTGCTCAAAATT 325
|||||
Db 730 ArgLysGluSerLeuSerLysIle 737

RESULT 11
E82193
conserved hypothetical protein VC1492 [imported] - *Vibrio cholerae* (strain N16961 ser
C:Species: *Vibrio cholerae*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: E82193
R:Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
Chadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: AB2035; MUID:20406833; PMID:10952301
A:Accession: E82193
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1613 <HEI>
A:Cross-References: GB:AE004352; NID:g9655990; PIDN:AAF94647.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1492
A:Map position: 1

Alignment Scores:
Pred. No.: 10.7 Length: 1613
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x E82193 (1-1613)

OY 254 AGAGTAAGCTGTAAAGCTGCA 231
|||||
Db 264 ArgValArgSerValLysLeuSer 271

RESULT 12
A60897
class I histocompatibility antigen H-2K(d) alpha chain, alternate splice form - mouse
C:Species: *Mus musculus* (house mouse)
C:Date: 04-Nov-1994 #sequence_revision 04-Nov-1994 #text_change 07-May-1999
C:Accession: A60897
R:Abu-Hadid, M.M.; Fujii, H.; Sood, A.K.
Mol. Immunol. 25, 739-749, 1988
A:Title: Identification of an alternatively spliced K(d) and the Qa-6(d) mRNAs by us1
A:Reference number: A60897; MUID:89039921; PMID:3141798
A:Accession: A60897
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-20 <ABU>

Alignment Scores:
Pred. No.: 192 Length: 20
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x A60897 (1-20)

OY 625 ACTTCGATTATCCCTGCC 605
|||||
Db 3 ThrSerAspLeuSerLeuPro 9

RESULT 13
S58771
NADH2 dehydrogenase (ubiquinone) (bc 1.6.5.3) chain 4L - *Pythium oligandrum* mitochond
C:Species: *Pythium oligandrum*
C:Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 03-Jun-2002
C:Accession: S58771
R:Martin, F.N.
Curr. Genet. 28, 225-234, 1995

A:Title: Linear mitochondrial genome organization in vivo in the genus *Pythium*.
A:Reference number: S58771; MUID:96071191; PMID:8529268
A:Accession: S58771
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-31 <MAR>
A:Cross-references: EMBL:U28355
C:Genetics:
A:Gene: nad4L
A:Genome: mitochondrion
A:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Alignment Scores:
Pred. No.: 184 Length: 31
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x S58771 (1-31)

OY 114 TATAAATAAAGTACTCTCA 94
|||||
DB 24 TyrlsTlleAsnLeuLeuSer 30

RESULT 14
E64453
hypothetical protein MJ1230 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: E64453
R:Butt, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
Ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese,
C
A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: E64453
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Residues: 1-76 <BDL>
A:Molecule type: DNA
A:Cross-references: GB:U67563; GB:L77117; NID:g22826379; PIDN:AAB99236.1; PID:g1591861; T
C:Genetics:
A:Map position: REV1173464-1173234
A:Start codon: TTG

Alignment Scores:
Pred. No.: 168 Length: 76
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x E64453 (1-76)

OY 267 AAGTCATGTCAGAGAGTAAG 247
|||||
DB 58 LysSerIleValLysSerLys 64

RESULT 15
T42310
hypothetical protein - phage SPPI
C:Species: phage SPPI
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
R:Alonso, J.C.; Luder, G.; Stiege, A.C.; Chai, S.; Weise, F.; Trautner, T.A.
Gene 204, 201-212, 1997
A:Title: The complete nucleotide sequence and functional organization of *Bacillus subtilis*
A:Reference number: 222137; MUID:98094274; PMID:9434185

A:Accession: T42310
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-91 <ALO>
A:Cross-references: EMBL:X97918; PIDN:CAA6517.1

Alignment Scores:
Pred. No.: 165 Length: 91
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x T42310 (1-91)

OY 342 GAGTCACGTGCAAAATTCAT 322
|||||
DB 41 GluSerLeuSerLysIleHis 47

RESULT 16
A02196
MHC class I histocompatibility antigen alpha chain (version 2) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 23-Jul-1999
C:Accession: A02196
R:Kress, M.; Giaros, D.; Khoury, G.; Jay, G.
Nature 306, 602-604, 1983
A:Title: Alternative RNA splicing in expression of the H-2K gene.
A:Reference number: A02196; MUID:84068207; PMID:6689056
A:Accession: A02196
A:Molecule type: mRNA
A:Residues: 1-102 <KR>
A:Cross-references: GB:K01762; NID:g199548; PIDN:AAA39655.1; PID:g387470
A:Experimental source: strain SWR, clone pH13
A>Note: this carboxyl-terminal fragment was translated from clone pH13, one of two R.
C:Genetics:
A:Insertions: 102/L
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: glycoprotein; heterodimer; transmembrane protein
F:78-102/Domain: Intracellular <INT>

Alignment Scores:
Pred. No.: 163 Length: 102
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x A02196 (1-102)

OY 625 ACTCTGATTATGCCGCGCC 605
|||||
DB 92 ThrSerAspLeuSerLeuPro 98

RESULT 17
A46264
thioredoxin 1 - slime mold (*Dictyostelium discoideum*)
C:Species: *Dictyostelium discoideum*
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A46264
R:Wetterauer, B.; Jacquot, J.P.; Veron, M.
J. Biol. Chem. 267, 9895-9904, 1992
A:Title: Thioredoxins from *Dictyostelium discoideum* are a developmentally regulated,
A:Reference number: A46264; MUID:92250653; PMID:1577820
A:Accession: A46264
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-105 <WET>
A:Cross-references: GB:M91384; NID:g167928; PIDN:AAA33258.1; PID:g167929
C:Superfamily: thioredoxin; thioredoxin homology
C:Keywords: redox-active disulfide

F;9-92/Domain: thioredoxin homology <THR>
F;32-35/Disulfide bonds: redox-active #status predicted

Alignment Scores:

Pred. No.:	163	Length:	105
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	1	Gaps:	0

US-09-817-318-1 (1-780) x A46264 (1-105)

OY 686 CTCAGTATGATTTATCACC 666

Db 45 LeuserAsngluPheIleThr 51

RESULT 18

A02197

MHC class I histocompatibility antigen alpha chain (version 1) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 23-Jul-1999

C:Accession: A02197

R:Kress, M.; Glaros, D.; Khoury, G.; Jay, G.

Nature 306, 602-604, 1983

A:Title: Alternative RNA splicing in expression of the H-2K gene.

A:Reference number: A02196; MUID:84068207; PMID:6689056

A:Accession: A02197

A:Molecule type: mRNA

A:Residues: 1111 <KRE>

A:Cross-references: GB:X00172; GB:X01445; NID:951328; PIDN:CAA24997.1; PID:951329

A:Experimental source: strain SMR, clone PH8

A>Note: the authors translated the codon CCG for residue 107 as Ser

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

C:Keywords: glycoprotein; heterodimer; transmembrane protein

F;78-111/Domain: Intracellular <INT>

Alignment Scores:

Pred. No.:	162	Length:	111
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	2	Gaps:	0

US-09-817-318-1 (1-780) x A02197 (1-111)

OY 625 ACTTCGATTTATCCCGCCC 605

Db 92 ThrSerAspLeuSerLeuPro 98

RESULT 19

I48639

neurotoxin homolog - mouse

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 23-Jul-1999

C:Accession: I48639

R:Fleming, T.J.; O'Nuigin, C.; Malek, T.R.

J. Immunol. 150, 5379-5390, 1993

A:Title: Characterization of two novel Ly-6 genes. Protein sequence and potential structure

A:Reference number: I48639; MUID:93294293; PMID:8515066

A:Accession: I48639

A:Molecule type: DNA

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-134 <RES>

C:Cross-references: EMBL:X70922; NID:9394727; PIDN:CAA50270.1; PID:9817978

C:Genetics:

A:Gene: Ly-6F.1

A:Introns: 24/1; 63/1

C:Superfamily: Ly-6 antigen; Ly-6 homology

Alignment Scores:

Pred. No.:	159	Length:	134
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.86%	Indels:	0
DB:	2	Gaps:	0

US-09-817-318-1 (1-780) x I48639 (1-134)

OY 701 GTTTAATTCCTCGTTGTA 721

Db 11 ValLeuIleLeuValVal 17

RESULT 20

T13311

hypothetical protein 22 - Streptococcus phage phi-O1205

C:Species: Streptococcus phage phi-O1205

C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 18-Aug-2000

C:Accession: T13311

R:Stanley, E.; Fitzgerald, G.F.; Le Marrec, C.; Fayard, B.; van Sinderen, D.

Microbiology 143, 3417-3429, 1997

A:Title: Sequence analysis and characterization of phi O1205, a temperate bacteriophage

A:Reference number: 217654; MUID:98048466; PMID:9387220

A:Accession: T13311

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-146 <STAS>

A:Cross-references: EMBL:U08974; NID:92444080; PID:92444102; PIDN:AACT9538.1

A:Experimental source: host Streptococcus thermophilus strain CNR21205

C:Superfamily: Streptococcus phage phi-O1205 hypothetical protein 22

Alignment Scores:

Pred. No.:	158	Length:	146
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.86%	Indels:	0
DB:	2	Gaps:	0

US-09-817-318-1 (1-780) x T13311 (1-146)

OY 295 GTTACAAAGATAGTCAG 315

Db 34 ValTyrIlySarGlyLeValLys 40

RESULT 21

T18586

hypothetical protein 6R55.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C:Accession: T18586

R:Wallis, J.

submitted to the EMBL Data Library, August 1998

A:Reference number: Z18992

A:Accession: T18586

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-147 <WIL>

A:Cross-references: EMBL:AL031222; PIDN:CAA20210.1; GSPDB:GN00028; CESP:6R55.2

A:Experimental source: clone 6R55

C:Genetics:

A:Gene: CESP:6R55.2

A:Map position: X

A:Introns: 50/2; 119/1

C:Superfamily: Caenorhabditis elegans hypothetical protein 6R55.2

Alignment Scores:

Pred. No.:	157	Length:	147
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.86%	Indels:	0
DB:	2	Gaps:	0

US-09-817-318-1 (1-780) x T18586 (1-147)

OY 415 TTTTCTCTGAGTTCCTTT 435

Db 114 PheserServalSerSerPhe 120

RESULT 22

G87632 hypothetical protein CC3097 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: G87632

R:Neuman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: AB7249; MUID:21173698; PMID:11259647

A:Accession: G87632

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1153 <STO>

A:Cross-references: GB:AE005673; NID:g13424753; PIDN:AAK25059.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC3097

Alignment Scores:

Pred. No.: 157 Length: 153

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.86% Indels: 0

DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x G87632 (1-153)

OY 227 TGTGTGACGCTTACAGACC 247

Db 144 CysValThrAlaLeuGlnThr 150

RESULT 23

158201 MHC class I antigen - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 21-Jan-2000

C:Accession: 158201

R:Jalanne, J.

Nucleic Acids Res. 10, 1039-1049, 1982

A:Title: Comparison of nucleotide sequences of mRNAs belonging to the mouse h-2 multigen

A:Reference number: 158201; MUID:82150234; PMID:6278432

A:Accession: 158201

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-164 <RES>

A:Cross-references: GB:J00395; NID:g199338; PIDN:AAA39579.1; PID:g387446

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

F,13-78/Domain: Immunoglobulin homology <IMM>

Alignment Scores:

Pred. No.: 156 Length: 164

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.92% Indels: 0

DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x 158201 (1-164)

OY 625 ACTTCTGATTATCCCTGCC 605

Db 145 ThrSerAspSerLeuPro 151

RESULT 24

D90066

hypothetical protein truncated-SA [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: D90066

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; (

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, I

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: D90066

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-168 <KUR>

A:Cross-references: GB:BA000018; PID:g13702553; PIDN:BA843694.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: truncated-SA

Alignment Scores:

Pred. No.: 155 Length: 168

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.86% Indels: 0

DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x D90066 (1-168)

OY 704 TTAATTCCTCGTTGTAATA 724

Db 10 LeuIleuLeuValIle 16

RESULT 25

A35145

transcription repressor of sporulation, separation and degradation paia - Bacillus suti

C:Species: Bacillus subtilis

C:Date: 03-Aug-1990 #sequence_revision 03-Aug-1990 #text_change 20-Jun-2000

C:Accession: A35145; E69671

R:Honjo, M.; Nakayama, A.; Fukazawa, K.; Kawamura, K.; Ando, K.; Horii, M.; Furutani,

J. Bacteriol. 172, 1783-1790, 1990

A:Title: A novel Bacillus subtilis gene involved in negative control of sporulation

A:Reference number: A35145; MUID:90202692; PMID:2108124

A:Accession: A35145

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-172 <HON>

A:Cross-references: GB:M36471; NID:g143283; PIDN:AAA22638.1; PID:g143284

A:Experimental source: strain M168

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Be

C.; Bron, S.; Brouillet, S.; Buschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Ga

tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hult, I

Koelter, P.; Koningsstein, G.; Krogh, S.; Kuno, M.; Kurita, K.; Lapidis, A.; Lardinc

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Ma

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portier

Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scani

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sc

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Toiguchi, A.; Tosato, V.; Uchiya

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshid

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: E69671

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-26, 'T', 28-172 <KUN>

A:Cross-references: GB:J29120; GB:AL009126; NID:g2635613; PIDN:CAB15205.1; PID:g26357

A:Experimental source: strain 168

C:Genetics:
A:Gene: pailA
C:Superfamily: Bacillus subtilis transcription regulator
C:Keywords: DNA binding; transcription regulation

Alignment Scores:
Pred. No.: 155 Length: 172
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x A35145 (1-172)

OY 504 TCTTCCAAAGCATGGCGT 484

Db 101 SerpneGlnYshSgLYleu 107

RESULT 26

JT0302
DNA-directed RNA polymerase (EC 2.7.7.6) delta chain rpoE - Bacillus subtilis

C:Species: Bacillus subtilis
C>Date: 31-Mar-1992 #sequence_revision 06-Feb-1995 #text_change 16-Jun-2000

C:Accession: JT0302; H32354; S55422; H69698

R: Lampe, M.; Binnie, C.; Schmidt, R.; Losick, R.

Gene 67, 13-19, 1988

A:Title: Cloned gene encoding the delta subunit of Bacillus subtilis RNA polymerase.

A:Reference number: JT0302; MUID:88329737; PMID:2843435

A:Accession: JT0302

A:Molecule type: DNA

A:Residues: 1-173 <L&M>

A:Cross-References: GB:M21677; NID:9143455; PIDN:AAA22710.1; PID:9143456

A:Note: Part of this sequence, including the amino end of the mature protein, was confir

R: Trach, K.; Chapman, J.W.; Piggett, P.; Lecocq, D.; Hoch, J.A.

J. Bacteriol. 170, 4194-4208, 1988

A:Title: Complete sequence and transcriptional analysis of the spoD' region of the Bacil

A:Reference number: A91883; MUID:88314920; PMID:2457578

A:Accession: H32354

A:Molecule type: DNA

A:Residues: 143-173 <TRA>

A:Cross-References: GB:M22039; NID:9460910; PIDN:AAA16800.1; PID:9143596

R: Glaser, P.; Danchin, A.

submitted to the EMBL Data Library, May 1995

A:Description: Cloning and sequencing of the Bacillus subtilis chromosomal region from 3

A:Reference number: S55414

A:Accession: S55422

A:Molecule type: DNA

A:Residues: 1-173 <GLA>

A:Cross-References: EMBL:249782; NID:9853752; PIDN:CAA89869.1; PID:9853761

R: Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter

A.: Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.: Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall

iech, J.; Hatwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.

Koelet, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kunita, K.; Lapidus, A.; Lardinois,

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Ser

deuch, M.; Tanaka, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: H69698

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-173 <KUN>

A:Cross-References: GB:299123; GB:AL009126; NID:92636240; PIDN:CA815744.1; PID:92636253

A:Experimental source: strain 168

C:Comment: This protein binds to the RNA polymerase core enzyme with or without a sigma

ruation.

C:Genetics:

A:Gene: rpoE

A:Start codon: TTG

C:Superfamily: DNA-directed RNA polymerase delta chain

C:Keywords: nucleotidyltransferase

F:2-173/Product: DNA-directed RNA polymerase delta chain #status experimental <MAT>

Alignment Scores:
Pred. No.: 155 Length: 173
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 1 Gaps: 0

US-09-817-318-1 (1-780) x JT0302 (1-173)

OY 760 GTAAGGCAAAAAAAAAA 780

Db 95 VallysAlalysLysLysLys 101

RESULT 27

A39022
acetylcholinesterase (EC 3.1.1.7), venom - Asian cobra (fragments)

C:Species: Naja naja oxiana (Asian cobra, Oxy cobra)

C>Date: 06-Dec-1991 #sequence_revision 06-Dec-1991 #text_change 14-Nov-1997

C:Accession: A39022

R: Weise, C.; Kretzenkamp, H.J.; Raba, R.; Aaviksaar, A.; Hucho, F.

J. Protein Chem. 9, 53-57, 1990

A:Title: The active site and partial sequence of cobra venom acetylcholinesterase.

A:Reference number: A39022; MUID:90253585; PMID:2340076

A:Accession: A39022

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-173 <ME>

C:Superfamily: cholinesterase; cholinesterase homology

C:Keywords: carboxylic ester hydrolase; glycoprotein

Alignment Scores:
Pred. No.: 155 Length: 173
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x A39022 (1-173)

OY 327 TTTTGACAGTGACTCTTTC 347

Db 59 PheLeuThrValThrLeuPhe 65

RESULT 28

S45548
hypothetical protein ypuF - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 28-Jul-2000

C:Accession: S45548; E69942

R: Sorokin, A.; Zumbstein, E.; Azevedo, V.; Ehrlich, S.D.; Serror, P.

submitted to the EMBL Data Library, November 1993

A:Reference number: S45533

A:Accession: S45548

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-174 <SOR>

A:Cross-References: EMBL:L09228; NID:9410114; PIDN:AAA67486.1; PID:9410130

R: Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber

C.: Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;

A.: Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal

iech, J.; Hatwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M

Koelter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetille, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadle, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeder, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Serot, akuchl, M.; Tamakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchilyama, T.; Winters, P.; Wipit, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:96044033; PMID:9384377
A:Accession: E69942
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-174 <KUN>
A:Cross-references: GB:299116; GB:AL009126; NID:g2634723; PIDN:CAB14255.1; PID:e1185592;
A:Experimental source: strain 168
C:Genetics:
A:Gene: ypuF
A:Start codon: TTG
C:Superfamily: *Bacillus subtilis* hypothetical protein ypuF

Alignment Scores:
Pred. No.: 155 Length: 174
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x S45548 (1-174)

OY 603 TTATCTAATCTTTTGTATT 583
|||||
Db 141 LeuSerAsnIlePheLeuVal 147

RESULT 29
D75514
hypothetical protein - *Deinococcus radiodurans* (strain R1)
C:Species: *Deinococcus radiodurans*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: D75514
R:White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: D75514
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-175 <WHI>
A:Cross-references: GB:AE001907; GB:AE000513; NID:g6456162; PIDN:AAF10064.1; PID:g645817
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0481
A:Map position: 1

Alignment Scores:
Pred. No.: 155 Length: 175
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x D75514 (1-175)

OY 701 GTTTAATTCCTCGTGTGA 721
|||||
Db 4 ValIleuIleuValVal 10

RESULT 30
S14981

extensin class I (clone w1-8 L) - tomato (fragment)
C:Species: *Lycopersicon esculentum* (tomato)
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 11-Jan-2000
C:Accession: S14981
R:Showalter, A.M.; Zhou, J.; Rumeau, D.; Worst, S.G.; Varner, J.E.
Plant Mol. Biol. 16, 547-565, 1991
A:Title: Tomato extensin and extensin-like cDNAs: structure and expression in respon
A:Reference number: S14970; MUID:91329690; PMID:1714316
A:Accession: S14981
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-199 <SHO>
A:Cross-references: EMBL:X55692
A:Experimental source: cv. UC82B
C:Superfamily: hydroxyproline-rich glycoprotein
C:Keywords: cell wall; glycoprotein; hydroxyproline

Alignment Scores:
Pred. No.: 153 Length: 199
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x S14981 (1-199)

OY 211 TTACCAATTCCTCCCAAAAA 191
|||||
Db 23 LeuThrIleSerSerLysLys 29

RESULT 31
T31585
hypothetical protein Y48C3A.c - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
C:Accession: T31585
R:Wallis, J.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z21046
A:Accession: T31585
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-206 <WIL>
A:Cross-references: EMBL:AL117203; PIDN:CAB55100.1; CESP:Y48C3A.c
A:Experimental source: clone Y48C3A
C:Genetics:
A:Gene: CESP:Y48C3A.c
A:Introns: 47/3; 89/2; 158/2
C:Superfamily: *Caenorhabditis elegans* hypothetical protein Y48C3A.c

Alignment Scores:
Pred. No.: 152 Length: 206
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.86% Indels: 0
DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x T31585 (1-206)

OY 701 GTTTAATTCCTCGTGTGA 721
|||||
Db 30 ValIleuIleuValVal 36

RESULT 32
T17228
hypothetical protein DKFZP434D146.1 - human (fragment)
C:Species: *Homo sapiens* (man)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T17228
R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999

A:Reference number: 218722
 A:Accession: T17228
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-211 <KOE>
 A:Cross-references: EMBL:AL117429
 A:Experimental source: adult testis; clone DKFZp434D146
 C:Genetics:
 A:Note: DKFZp434D146.1

Alignment Scores:
 Pred. No.: 152 Length: 211
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.86% Indels: 0
 DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x T17228 (1-211)

OY 526 CTATTTCATGCTGCTACTT 546
 |||||||
 Db 84 LeuphetThrLeuphetValleu 90

RESULT 33

45997
 T45997
 Hypothetical protein F9D24.280 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 02-Sep-2000
 C:Accession: T45997
 R:D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.;
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: 223011
 A:Accession: T45997
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-219 <DAN>
 A:Cross-references: EMBL:AL137081
 A:Experimental source: cultivar Columbia; BAC clone F9D24
 C:Genetics:
 A:Map position: 3
 A:Introns: 85/3
 A:Note: F9D24.280
 C:Superfamily: Arabidopsis thaliana hypothetical protein F9D24.210

Alignment Scores:
 Pred. No.: 151 Length: 219
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.86% Indels: 0
 DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x T45997 (1-219)

OY 760 GTAAAGCAAAAAAAAAAAAAA 780
 |||||||
 Db 164 VallysAlaLysLysLysLys 170

RESULT 34

559134
 S59134
 Probable zinc finger transcription activator Rep2 - fission yeast (Schizosaccharomyces f.
 C:Species: Schizosaccharomyces pombe
 C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 31-Jan-2000
 C:Accession: S59134; T40130
 R:Nakashima, N.; Tanaka, K.; Sturm, S.; Okayama, H.
 EMO J. 14, 4794-4802, 1995
 A:Title: Fission yeast Rep2 is a putative transcriptional activator subunit for the cell
 A:Reference number: S59134; MID:96030785; PMID:7588609
 A:Accession: S59134
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-219 <NAK>

A:Cross-references: EMBL:X91044; NID:9975711; PIDN:CAA62504.1; PID:9975712
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.
 submitted to the EMBL Data Library, June 1997
 A:Reference number: 221907
 A:Accession: T40130
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-219 <WOO>
 A:Cross-references: EMBL:Z97211; PIDN:CA10158.1; GSPDB:GN00067; SPDB:SPBC2F12.11C
 A:Experimental source: strain 972h-; cosmid cCF12
 C:Genetics:
 A:Gene: SPBC2F12.11C
 A:Map position: 2

Alignment Scores:
 Pred. No.: 151 Length: 219
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.86% Indels: 0
 DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x S59134 (1-219)

OY 418 TCTCTGTGAGTCTCCTTCC 438
 |||||||
 Db 42 SerServalSerPheAla 48

RESULT 35

540931
 S40931
 Hypothetical protein ZK1098.9 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
 C:Accession: S40931
 R:Thomas, K.
 submitted to the EMBL Data Library, February 1992
 A:Reference number: S40923
 A:Accession: S40931
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-220 <THO>
 A:Cross-references: EMBL:Z22176; NID:9297978; PID:9297987
 C:Genetics:
 A:Introns: 48/1; 86/3; 111/3; 145/3; 188/3

Alignment Scores:
 Pred. No.: 151 Length: 220
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.86% Indels: 0
 DB: 2 Gaps: 0

US-09-817-318-1 (1-780) x S40931 (1-220)

OY 546 TGCAATTGTATGCTAGATGTT 566
 |||||||
 Db 183 CysIleValleuAspVal 189

RESULT 36

AI0781
 AI0781
 GTP cyclonucleotidylase I (EC 3.5.4.16) [imported] - Salmonella enterica subsp. enterica s
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AI0781
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
 et al.; Connerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
 S.; Moutle, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
 A:Reference number: AB0502; PMID:11677608

A:Accession: A10781
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-222 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD02575.1; PID:gl6503433; GSPDB:GN00176
A:Gene: STY2427
C:Superfamily: GTP cyclohydrolase I
C:Keywords: hydrolase

Alignment Scores:

Pred. NO.:	151	Length:	222
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.86%	Indels:	0
DB:	2	Gaps:	0

US-09-817-318-1 (1-780) x A10781 (1-222)

OY 233 ACAGCTTACAGACCTTACTC 253
|||||
Db 161 ThrAlaLeuGlnThrLeu 167

RESULT 37
D89836
hypothetical protein SA0614 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: D89836
R:Kurodo, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratazu, K.;
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: D89836
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-224 <KUR>
A:Cross-references: GB:BA000018; PID:gl3700550; PIDN:BA041847.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA0614
C:Superfamily: ompR protein; response regulator homology

Alignment Scores:

Pred. NO.:	151	Length:	224
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	2	Gaps:	0

US-09-817-318-1 (1-780) x D89836 (1-224)

OY 542 ACAGTCATGTAATAGATTG 522
|||||
Db 193 ThrValAsnValAsnArgLeu 199

RESULT 38
G70000
two-component response regulator [ytsH] homolog ytsA - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: G70000
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emerson, P.T.; Estlin, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koelter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; May, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Porter, Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl, A:Authors: Schleich, S.; Schroeter, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Sakeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terstie, P.; Tognoni, A.; Tosato, V.; Uchida, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: G70000
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-231 <STO>
A:Cross-references: GB:D99119; GB:AL009126; NID:g2635411; PIDN:CAB15018.1; PID:g2635
A:Experimental source: strain 168
C:Genetics:
A:Gene: ytsA
C:Superfamily: ompR protein; response regulator homology
C:Keywords: phosphoprotein
F:4-112/Domain: response regulator homology <RR>
F:52/Binding site: phosphate (asp) (covalent) #status predicted

Alignment Scores:

Pred. NO.:	151	Length:	231
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	2	Gaps:	0

US-09-817-318-1 (1-780) x G70000 (1-231)

OY 542 ACAGTCATGTAATAGATTG 522
|||||
Db 194 ThrValAsnValAsnArgLeu 200

RESULT 39
G84138
two-component response regulator BH3911 [imported] - Bacillus halodurans (strain C-1.
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: G84138
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: G84138
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-231 <STO>
A:Cross-references: GB:AP001520; GB:BA000004; NID:gl0176401; PIDN:BA07630.1; GSPDB: A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3911
C:Superfamily: ompR protein; response regulator homology

Alignment Scores:

Pred. NO.:	151	Length:	231
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.92%	Indels:	0
DB:	2	Gaps:	0

US-09-817-318-1 (1-780) x G84138 (1-231)

OY 542 ACAGTCATGTAATAGATTG 522
|||||
Db 194 ThrValAsnValAsnArgLeu 200

RESULT 40
S43513
hemoglobin linker chain 2 - polychaete (Mearthes diversicolor) (fragment)

```
C:Species: Neanthes diversicolor
C:Date: 19-Mar-1997 #sequence_revision 05-Dec-1997 #text_change 18-Aug-2000
C:Accession: S43513
R:Suzuki, T.; Ohla, T.; Yuasa, H.-J.; Takagi, T.
Biochim. Biophys. Acta 1217, 291-296, 1994
A:Title: The giant extracellular hemoglobin from the polychaete Neanthes diversicolor.
A:Reference number: S43513; M01D:94198291; PMID:8148374
A:Accession: S43513
A:Molecule type: mRNA
A:Residues: 1-235 <SUZ>
A:Cross-references: EMBL:D58413; NID:g893395; PIDN:BA09580.1; PID:d1010223; PID:g893395
C:Function: essential for assembly of heme containing subunits of extracellular giant
A:Note: annelid-like extracellular hemoglobins differ from all other vertebrate or non-
C:Superfamily: hemoglobin linker chain; LDL receptor ligand-binding repeat homology
F:68-106/Domain: LDL receptor ligand-binding repeat homology <LDL2>

Alignment Scores:
Pred. No.: 150 Length: 235
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x S43513 (1-235)
QY 470 TTGGGTGATTTCAGAGACTG 450
|||||
Db 50 Leuclyaspheargicluen 56

RESULT 41
D70032
two-component response regulator [YvcQ] homolog yvcP - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: D70032
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berlet
C: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A: Ehlich, S.D.; Emmerson, P.T.; Eutian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gall
leeh, J.; Harwood, C.R.; Henalt, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetler, P.; Koningstein, G.; Kiroh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
Y. M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sdaite, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schreter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Ser
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; M01D:98044033; PMID:9384377
A:Accession: D70032
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-237 <KUN>
A:Cross-references: GB:299121; GB:AL009126; NID:g2635827; PIDN:CAB15477.1; PID:g2635985
A:Experimental source: strain 168
C:Genetics:
A:Gene: yvcP
C:Superfamily: ompr protein; response regulator homology
C:Keywords: phosphoprotein
F:4-112/Domain: response regulator homology <RRH>
F:52/Binding site: phosphate (Asp) (covalent) #status predicted

Alignment Scores:
Pred. No.: 150 Length: 237
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x D70032 (1-237)
QY 542 ACAGTCATGAAATAGATG 522
|||||
Db 198 ThrValasNValasArgLeu 204

RESULT 42
A81422
probable RNA polymerase sigma factor for flagellar operon Cj061c [imported] - Campy
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: A81422
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chl
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Bar
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A:Reference number: A81250; M01D:20150912; PMID:10688204
A:Accession: A81422
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-238 <PAR>
A:Cross-references: GB:AL139074; GB:AL111168; NID:g6967505; PIDN:CAB72548.1; PID:g696
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: fliA; Cj0061c
C:Superfamily: transcription initiation factor sigma; transcription initiation facto

Alignment Scores:
Pred. No.: 150 Length: 238
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.92% Indels: 0
DB: Gaps: 0

US-09-817-318-1 (1-780) x A81422 (1-238)
QY 330 AAATTCATGAGTTCGAC 310
|||||
Db 182 LysIleHisGluValLeuasp 188

RESULT 43
A38379
phosducin, retinal - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 28-Jun-1991 #sequence_revision 14-Jul-1994 #text_change 04-Oct-1996
C:Accession: A38379; J02025; A38378
R:Lee, R.H.; Fowler, A.; McGinnis, J.F.; Lolley, R.N.; Craft, C.M.
J. Biol. Chem. 265, 15867-15873, 1990
A:Title: Amino acid and cDNA sequence of bovine phosducin, a soluble phosphoprotein f
A:Reference number: A38379; M01D:90368806; PMID:2203790
A:Accession: A38379
A:Molecule type: protein
A:Residues: 1-245 <LEE>
R:Dee, T.; Nakabayashi, H.; Tamada, H.; Takagi, T.; Sakurai, S.; Yamaki, K.; Shinoha
Gene 91, 209-215, 1990
A:Title: Analysis of the human, bovine and rat 33-kDa proteins and cDNA in retina and
A:Reference number: J02025; M01D:91007277; PMID:2210381
A:Accession: J02025
A:Molecule type: mRNA
A:Residues: 1-43; 'P', 45-238; 'K', 240-245 <ABE>
A:Cross-references: GB:M33529
R:Lee, R.H.; Brown, B.M.; Lolley, R.N.
J. Biol. Chem. 265, 15860-15865, 1990
A:Title: Protein kinase A phosphorylates retinal phosducin on serine 73 in situ.
A:Reference number: A38378; M01D:90368805; PMID:2394752
A:Accession: A38378
A:Molecule type: protein
A:Residues: 54-58; 'X', 60-61; 'X', 63-64; 'XX', 67-70; 'X', 72-88; 'X', 90; 'X', 92-93; 102-114, '
C:Superfamily: phosducin
C:Keywords: glycoprotein; phosphoprotein; photoreceptor; retina
F:8,58,238/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status pre
```

F:45/69/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
F:73/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status experimental
F:152/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:

Alignment Scores:	150	Length:	245
Pred. No.:	7.00	Matches:	7
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	2.86%	Gaps:	0
DB:	1		

US-09-817-318-1 (1-780) x A38379 (1-245)

QY 679 TTACTGACATTCATATCA 699

DB 197 LeuLeuSerAsnPhelLeu 203

RESULT 44

G81297

hypothetical protein Cj1507c [imported] - Campylobacter jejuni (strain NCTC 11168)

C:Species: Campylobacter jejuni

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002

C:Accession: G81297

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chilling

C.W.; Oual, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: G81297

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-245 <PAR>

A:Cross-references: GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAB73928.1; PID:g696893

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: Cj1507c

C:Superfamily: Campylobacter jejuni hypothetical protein Cj1507c

Alignment Scores:

Alignment Scores:	150	Length:	245
Pred. No.:	7.00	Matches:	7
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	2.86%	Gaps:	0
DB:	2		

US-09-817-318-1 (1-780) x G81297 (1-245)

QY 297 TTACAAAGATGCAAGAA 317

DB 187 LeuGlnLysAsnSergInGln 193

RESULT 45

T17397

vr1Q protein - Dichelobacter nodosus

C:Species: Dichelobacter nodosus

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C:Accession: T17397

R:Billington, S.J.; Huggins, A.S.; Johanesen, P.A.; Crellin, P.K.; Cheung, J.K.; Katz, M

Infect. Immun. 67, 1277-1286, 1999

A:Title: Complete nucleotide sequence of the 27-kilobase virulence related locus (vrl) C

A:Reference number: Z18734; MUID:99150261; PMID:10024571

A:Accession: T17397

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-251 <BIL>

A:Cross-references: EMBL:U02046; NID:g3493323; PID:g2317810; PIDN:AAC33386.1

A:Experimental source: strain A198

C:Superfamily: Dichelobacter nodosus vrlQ protein

Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.86%	Indels:	0
DB:	2	Gaps:	0

US-09-817-318-1 (1-780) x T17397 (1-251)

QY 312 CAAGACTTCATGAATTTTG 332

DB 37 GlnGlnLeuHisGlnPhelLeu 43

Search completed: February 4, 2003, 07:23:29
Job time : 53 secs

